Repair by ellipsis in BCS*

Anna Grabovac University of Maryland

1 Introduction

In most contexts, sluicing (TP-ellipsis) in Bosnian/Croatian/Serbian (BCS) behaves straightforwardly. As shown in (1), the case realized on the *wh*-remnant must match the case on its correlate in the antecedent. This apparent case connectivity is observed in both structural (NOM, ACC) and inherent (GEN, DAT, INSTR) case environments¹.

```
(1) a.
        Vid(j)eli su
                           nekoga,
                                         ali ne znam
                 AUX.3PL someone.ACC but not I.know
         saw
                      / *tko
           {koga
                                  \{ su \ vid(j)eli \}.
            who.ACC / who.NOM
         'They saw someone, but I don't know who.'
       Približila
                    se
                          nekome,
                                        ali ne znam
         approached REFL someone.DAT but not I.know
                     / *koga
                                  / *tko
                                              } [<del>se približila</del>].
            who.DAT / who.ACC / who.NOM
         'She approached someone, but I don't know who.'
```

In this respect, BCS resembles many other languages. However, the overall account of BCS becomes more complicated when we consider examples containing numeral quantifiers. Following the literature, I refer to such examples using 'genitive of quantification' or GQ, which captures the observation that the numeral licenses genitive case on its complement (Bošković 2006; Stjepanović 2012). Interestingly, GQ examples do not display the same case connectivity effects as non-GQ examples, such as those in (1). A successful analysis must therefore maintain the strictness needed for the majority of BCS examples while also allowing enough flexibility to accommodate the GQ examples.

^{*}For questions and discussion at various stages of this work, I'd like to thank Klaus Abels, Jeff Lidz, Jason Merchant, Troy Messick, Masha Polinsky, Laura Stigliano, and Gary Thoms, as well as three anonymous reviewers and the attendees of CLS 59. I am also grateful to Milena Stajić, Jana Willer-Gold, and many anonymous consultants for help with the data. Any remaining errors are my own.

¹Glossing abbreviations: 3 = third person, ACC = accusative, AGR = agreement, AOR = aorist, AUX = auxiliary, COMP = complementizer, DAT = dative, DIR = direct, GEN = genitive, INSTR = instrumental, NOM = nominative, OBL = oblique, PL = plural, PROG = progressive, PST = past, QUE = question, REFL = reflexive, SG = singular

When a GQ construction occurs in a structural case environment, the genitive licensed by the numeral takes priority over the externally licensed case; the numeral itself is indeclinable. For example, (2) shows the GQ-construction in an accusative context, where genitive rather than accusative is realized.

(2) Koliko žirafa su vid(j)eli? how.many giraffes.GEN AUX.3PL saw 'How many giraffes did they see?'

An account of these examples must explain why GQ takes priority over structural case (see e.g., Franks 1995; Bošković 2006; Šarić 2014; Grabovac 2022), but in this respect, BCS is not unique among the Slavic languages. Unsurprisingly, related examples are grammatical under sluicing.

(3) Vid(j)eli su jedn-u zebr-u, ali ne znam saw AUX.3PL one-ACC zebra-ACC but not I.know koliko žirafa.
how.many giraffes.GEN
'They saw one zebra, but I don't know how many giraffes.'

However, inherent case contexts present a puzzle. GQ constructions are ungrammatical as complements of dative- and instrumental-case-assigning verbs, a unique characteristic of BCS. Nonetheless, as Stjepanović (2012) initially observed, such examples are grammatical under sluicing, exemplified in (4b). Note that for reasons of space, this paper focuses on dative contexts; instrumental examples are slightly more involved as they involve the preposition *sa* 'with' (see Stjepanović 2013; Grabovac 2022, ch.4 for further discussion).

- (4) a. *Koliko žirafa se približila? how.many giraffes.GEN REFL approached 'How many giraffes did she approach?'
 - b. *Približila se jedn-oj zebr-i, ali ne znam* approached REFL one-DAT zebra-DAT but not I.know *koliko žirafa*. how.many giraffes.GEN

'She approached one zebra, but I don't know how many giraffes.'

As it turns out, there is no grammatical pre-sluice for (4b). In addition to the syntactically identical option in (4a), the passive and cleft constructions in (5) are also ruled out.²

²The passive was emphatically ruled out by my consultants, with some even commenting that they could not formulate it. The cleft in (5b) is reportedly less bad than the passive, though inappropriate for the context (see §4.1 for discussion).

- (5) a. *Približila se jedn-oj zebr-i, ali ne znam approached REFL one-DAT zebra-DAT but not I.know koliko žirafa je približ-eno (od nje). how.many giraffes.GEN AUX.3SG approach-PASS by her
 - 'She approached one zebra, but I don't know how many giraffes were approached (by her).'
 - b. ??*Približila se jedn-oj zebr-i, ali ne znam* approached REFL one-DAT zebra-DAT but not I.know *koliko je to žirafa bilo.*

how.many AUX.3SG that giraffes.GEN was

'She approached one zebra, but I don't know how many giraffes it was.'

This pattern is not limited to sluicing. As shown in (6), the ungrammaticality of GQ/inherent case also disappears under VP-ellipsis.³

(6) *Žena se približila jedn-oj žiraf-i, ali ne znam* woman REFL approached one-DAT giraffe-DAT but not I.know

koliko se žirafa d(j)evojka (*približila). how.many REFL giraffes.GEN girl approached

'The woman approached one giraffe, but I don't know how many giraffes the girl did.'

Stjepanović (2013) further observes that gapping constructions show similar effects in dative contexts. Note that BCS higher numerals (≥ 5) behave the same as the numeral quantifier *koliko* that we see in (2)–(6).

(7) a. *On će prići jedn-oj zebr-i*, he will approach one-DAT zebra-DAT

a ona (će) pet žirafa. and she will five giraffes.GEN

'He will approach one zebra, and she (will) five giraffes.'

b. **On će prići jedn-oj zebr-i*, he will approach one-DAT zebra-DAT

> a ona će prići pet žirafa. and she will approach five giraffes.GEN

'He will approach one zebra, and she will approach five giraffes.'

³The past tense in BCS is formed with a copular auxiliary that reflects the features of the subject and a past participle. For verbs that obligatorily include reflexive se, the 3.SG auxiliary je is omitted in the formation of the past tense. Thus, while the embedded clause in (6) appears to lack any sort of verbal form, se in fact functions as a substitute for je.

Given the lack of acceptable pre-elliptical structures for GQ examples, I suggest that (4b), along with (6) and (7), are instances of repair by ellipsis. The question is, what specifically is being repaired?

2 The Inverse Inherent Case Filter

In an initial account of GO/inherent case examples under sluicing, Stiepanović (2012) suggests that the issue to be repaired stems from the inability of inherent case to be realized. Building off the proposal in Bošković 2008, Stjepanović adds a morphological requirement to the Inverse Inherent Case Filter (IICF): "Inherent case must be morphologically realized, if it can be [emphasis mine]. A verb lexically specified for inherent case must assign its morphological case feature to the NP that checks its abstract case feature against the verb in syntax" (2012:80). Crucially, Stjepanović's proposal is PF-based, in contrast to the LF-centered account in Bošković 2008. Bošković contends that ungrammaticality of examples such as (4a) is due to a violation of the Theta Criterion. Since the numeral quantifier is caseless, the inherent-case-assigning verb cannot check its case against it. This is problematic since an inherent-case-assigning verb is assumed to theta-mark its complement if and only if it can assign the inherent case (Chomsky 1986). Structural case is not associated with theta role assignment, so examples like (2) are not predicted to be problematic. However, Stjepanović (2012) points out that the LF representations of (4a) and the embedded clause in (4b) are the same. If (4a) is ungrammatical due to an LF-based IICF, then (4b) should be as well, contrary to what we see. Stjepanović therefore suggests that the proper account cannot come from an LF version of the IICF.

While I agree with the intuition behind Stjepanović's version of the IICF, it is too weak as is. The condition that 'inherent case must be morphologically realized, if it can be' is flexible enough to capture examples such as (8) from Russian, where an indeclinable noun can occur as the complement of an inherent-case-assigning verb.⁴

However, this formulation of the IICF is not strong enough to explain why examples such as (4a) are bad in BCS. In (4a), inherent case cannot be morphologically realized since the numeral is caseless. Consequently, the IICF states that it does not have to be realized. We are thus left without an account of why the example is ungrammatical. Although I agree with Stjepanović that the inability of inherent case to be realized is the source of the issue in BCS, the IICF cannot be the full story. The next section discusses how nominal concord provides insight into the particular nature of the case clash and subsequent repair by ellipsis.

⁴Note that Stjepanović 2013 omits this conditional statement in the IICF. However, the analysis must then say something more to accommodate examples such as (8). I take this as further evidence that the IICF cannot be the whole account.

3 Clues from concord

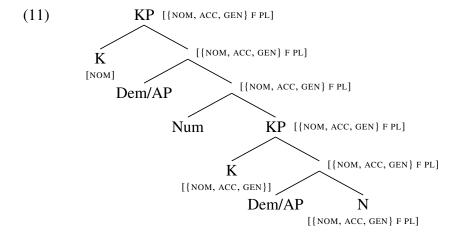
As mentioned earlier, BCS higher numerals (≥ 5) behave the same as the numeral quantifier *koliko* 'how many'. Grabovac 2022 provides a theory of the concord patterns displayed by these higher numeral constructions, and we will see that this theory provides insight into the nature of the repair in (4)–(7). In nominative, accusative, and genitive case environments, higher numeral constructions display the concord pattern in (9), where GEN.PL is realized above and below the numeral.

(9) *t-ih pet visok-ih žirafa* that-GEN.PL five tall-GEN.PL giraffe.GEN.F.PL 'those five tall giraffes'

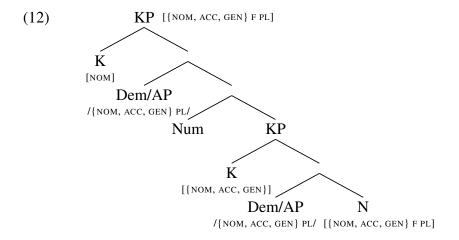
In dative and instrumental case contexts, on the other hand, higher numeral constructions are ungrammatical (Wechsler & Zlatic 2003; Bošković 2006; Šarić 2014), exemplified below. This is precisely the distribution displayed by *koliko* in (4).

(10) *Približila se t-ih pet visok-ih žirafa.
approached REFL that-GEN.PL five tall-GEN.PL giraffe.GEN.F.PL
'She approached those five tall giraffes.'

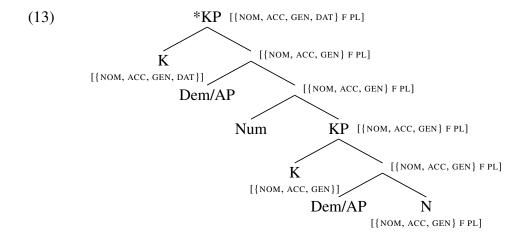
To get an idea of how the concord system works, let us proceed through a quick derivation of (9). Following Grabovac 2022, concord consists of syntactic and postsyntactic stages (see also Norris 2014; Ackema & Neeleman 2020). The assumed syntax of numeral constructions contains two extended projections (KP in the structures here), or two syntactic domains. The lower domain is consistently genitive {NOM, ACC, GEN} following the common observation that BCS numerals impose genitive case on their complements; the higher domain reflects the externally licensed case. In the syntactic stage of the derivation, features percolate upward. For simplicity of presentation, (11) depicts gender and number features percolating from the noun, but these could be introduced in their own projections higher in the structure with no change to the theory of concord. One key element of the derivation is case override. In (11), we see that the genitive case and ϕ -features licensed in the lower domain are able to percolate through the higher domain, ultimately reaching the higher KP. This extended percolation beyond the lower domain boundary is allowed because the numeral bears no competing features (reminiscent of 'relativized heads' in Di Sciullo & Williams 1987). The percolating set of {NOM, ACC, GEN} ultimately overrides {NOM} on the topmost KP given containment relations (see Caha 2009 for motivation behind the hierarchical decomposition of case).



The tree in (12) represents the post-syntactic mapping of the structure in (11), with the result of concord denoted by slashes. Note that the features on intermediate nodes have been omitted for convenience, but this may or may not be part of the actual derivation. In concord, features from the closest dominating node are realized on available terminals, resulting in GEN.PL throughout the construction.



Let us now consider what goes wrong in examples such as (10). The derivation begins as before, with feature percolation in the syntax. This time, however, the set of genitive features percolating from the lower domain is unable to override the set of dative features to reach the higher KP since {NOM, ACC, GEN, DAT} $\not\subset$ {NOM, ACC, GEN}. When concord spells out the most local set of dominating features, genitive rather than dative is realized throughout the construction. Grabovac 2022 argues for a condition whereby the dominating set of case features must be realized. Since there are no available terminals for the dominating dative case, the derivation is predicted to crash.



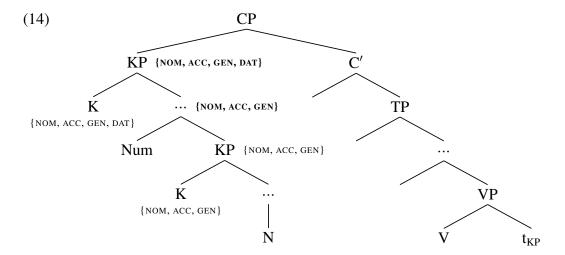
The issue in examples such as (10) stems from the inability of the dominating case features to be realized. This suggests that repair by ellipsis affects the features on KP.

4 Repair by ellipsis

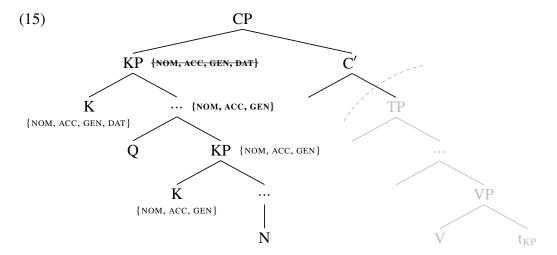
Building from the theory of concord outlined in the previous section, the nature of the repair relies on two main hypotheses. (i) Assuming a move-and-delete approach to ellipsis (Ross 1969; Merchant 2001; a.o.), case licensed on a lower copy of movement is also present on higher copies. (ii) If a case licenser is elided, the case features inherited by any elements preserved by movement may be deleted as well. Assuming the grammar seeks to avoid the deletion of features wherever possible, the latter hypothesis applies as a last resort. Thus, the deletion of case features in ellipsis functions as a repair. The following derivation of (4) illustrates how these hypotheses unfold.

4.1 A sluicing derivation

Given the theory of concord and the outcome of examples such as (13), we predict a case clash in the syntax, as shown in (14). The verb licenses dative on the higher KP, which is preserved on KP following movement. As usual, the numeral quantifier licenses genitive on the lower KP, which percolates through the construction. As above, {NOM, ACC, GEN} ceases to percolate since it cannot override {NOM, ACC, GEN, DAT} on KP due to a lack of containment.



In the previous derivation, we saw that when concord spells out the closest set of dominating features, genitive is realized throughout the numeral construction. This leaves no available terminals to realize the dative case licensed on the higher KP, and the derivation should crash. However, given hypothesis (ii), the dative features licensed by the verb can be deleted on KP when the entire TP is elided, thus eliminating the problem.



The last resort nature of hypothesis (ii) is key for non-GQ examples such as (1), which display case connectivity. Since feature deletion is avoided where possible (i.e., it is only triggered as a repair), case matching is preserved in the absence of a case clash.

While hypothesis (ii) is necessary for the current analysis, additional support for deletion of case features following deletion of the case licenser can be found in Brazilian Portuguese. Exceptive constructions such as (16) show that when the case licenser *para* is removed from the example, so is the oblique case it licenses (Pleshak 2023:5).

- (16) a. Ela deu presentes para todo mundo menos para mim. she gave presents to all world minus to I.OBL 'She gave presents to everyone but me.'
 - b. Ela deu presentes para todo mundo menos {eu /*mim}. she gave presents to all world minus I.DIR / I.OBL 'She gave presents to everyone but me.' (Brazilian Portuguese)

Independent support can be also found in certain examples of Turkish *pro*-drop. Although they do not reflect deletion of case following deletion of the case licenser, these examples provide a parallel with agreement morphology. Öztürk (2002:243) shows that overt pronouns can occur without agreement morphology in certain adjuncts, as in (17a). However, in the absence of an overt pronoun as in (17b), the interpretation changes to a third-person subject. Thus, when the overt pronoun is omitted, so is any interpretation of agreement.

- (17) a. *Ben konuş-ur-ken*, *o gül-üyor-du*. I talk-AOR-while s/he laugh-PROG-PST 'While I was talking, s/he was laughing.'
 - b. pro_i konuş-ur-ken, pro_i gül-üyor-du. talk-AOR-while laugh-PROG-PST 'While s/he was talking, s/he was laughing.' (*Turkish*)

It is also worth noting that the proposed analysis is compatible with the other potential pre-sluices in (5) and not only the syntactically identical option. In terms of the passive in (5a), last resort deletion and repair is predicted to take effect. Mc-Fadden (2004) points out that unlike in a structural-case context where the passive overrides the assignment of accusative case, inherent case is still typically preserved in the passive construction. Thus, the presence of GQ clashes with the passivized verb in the same way it clashes with the non-passivized form. Regarding the cleft in (5b), GQ takes precedence over structural case (as discussed in §3), so there is no case clash and thus no deletion of case features with ellipsis. However, Reeve (2012) reports that BCS cleft constructions are not felicitous in contrastive contexts, such as GQ examples where the quantified noun presents an alternative to the correlate. This accounts for the relatively poor rating of (5b)⁵.

4.2 Beyond sluicing

As discussed in §1, VP-ellipsis and gapping provide further evidence of repair when GQ constructions occur as the complements of dative-assigning verbs. The analysis outlined above for the sluicing examples is applicable. In VP ellipsis, the verb licenses inherent case on KP of the numeral construction, which is preserved when the numeral construction undergoes movement. The numeral quantifier also licenses genitive case internal to the numeral construction, which percolates but

⁵Non-GQ examples can occur with a cleft, which licenses nominative on the wh-phrase, but this cleft cannot be elided (Stjepanović 2012). The impossibility of eliding the cleft suggests that the analysis may require an identity condition on sluicing (see Saab 2010, a.o.).

fails to reach KP. Normally, this would result in a crash, but the offending case features are eliminated at PF when the verb is elided. Similarly, in gapping constructions, deletion of the case features licensed by the verb occurs as a last resort repair of the case clash.

5 Possible alternative: Ellipsis bleeds m-case

Thoms 2019 outlines an account very similar to the one proposed in this paper. The core of the proposal is that ellipsis bleeds morphological case; thus, ellipsis amounts to non insertion in morphology rather than PF-deletion. When TP is elided in sluicing, the morphological case features that would determine the case of the remnant are suppressed. In the absence of these features, the remnant can get case from a local licenser (such as the numeral quantifier in my examples), but in the absence of a local licenser (as in the case matching examples), the remnant enters into a syntactic dependency with the correlate. This means that the correlate's syntactic context determines the remnant's case. Whereas my proposal requires us to accept the possibility of deletion of case features under ellipsis, Thoms's proposal requires a multidominance structure to value the remnant's case (see also van Craenenbroeck 2017).

These slight differences in implementation bring about differences in the predictions made by our proposals. For Thoms, case on the remnant is externally conditioned by the correlate's context in the absence of a more local licenser. However, it is not always apparent that the case on the remnant aligns with this prediction. For example, Chamorro (Chung 2013, as cited in Vicente 2015:17) indicates that the case on the remnant does not correspond to that of the correlate. Moreover, there is no obvious local licenser.

(18) Ilek-ñiha na man-ma'à'ñao siha ni un tảotao, lao ti ma say-AGR COMP AGR-afraid they OBL a person but not AGR sảngan (*ni) hảyi.

say OBL who

'They said they were afraid of a certain person, but they didn't say who.'
(Chamorro)

This is not necessarily fatal for the analysis—examples such as (18) may prove amenable to the account Thoms outlines for Mongolian, which also displays case mismatches between the correlate and the remnant without an obvious local licenser. Moreover, any analysis of (18) will depend on the nature of the pre-sluice and the theory's construal of the identity between the elided material and antecedent material. As it stands, my proposal would also need to say something more to capture these examples, unless a motivation for repair can be found.

Some languages, such as Korean, contain an apparent case mismatch but clearly do not require the same repair proposed for BCS (see e.g., Kim 2012; Park 2014; Nykiel *et al.* 2023). As shown in (19), Korean allows a caseless fragment, even when the correlate realizes accusative (Nykiel *et al.* 2023:330). While this resembles deletion of case, as proposed in my analysis, Korean in fact allows case to be

dropped from non-nominative NPs (Kim 2015). Thus, the pre-elliptical structure may have been caseless in the first place.

(19) A: Mimi-ka mwues-ul masy-ess-ni? B: Cwusu. A: Mimi-NOM what-ACC drink-PST-QUE B: juice

A: 'What did Mimi drink?' B: 'Juice.' (Korean)

It could be possible to relax the last resort nature of hypothesis (ii), subject to language-specific tendencies, but this is not an ideal solution. Space constraints preclude a more detailed discussion, but hopefully this overview lays the groundwork for further dialogue and future research.

6 Outlook

This paper has examined elliptical repair of GQ constructions in BCS. While BCS typically exhibits standard case connectivity effects, GQ examples display a case mismatch between the correlate and the remnant. This alone is not problematic, but the overall picture becomes more complicated in inherent case environments, which have no grammatical pre-elliptical structure. I proposed that concord offers insights into how this repair should be analyzed. In particular, the derivation crashes when dominating case features cannot be realized, so elimination of the problematic case features at PF saves the derivation. Under a PF view of ellipsis, deletion of the case licenser allows the offending case features to be deleted as a last resort.

One prominent question in research on ellipsis concerns the nature of the identity between the elided material and the antecedent material. Those who accept the existence of syntactic structure at the ellipsis site can be divided into two groups: some argue for complete or near identity between the antecedent and the elided material (e.g., Ross 1969; Lasnik 2005), while others adopt a more nuanced view. For the latter group, the elided material need not be syntactically identical, but well-formed with respect to the antecedent, as well as semantically appropriate (e.g., Merchant 2001; Barros *et al.* 2014; Abels 2017). The BCS data are inconclusive on this front. Given the repair story adopted in this paper, any of the possible pre elliptical structures, syntactically identical or not, are compatible with the analysis. This merely amplifies the question of how to properly formulate the identity condition.

References

Abels, K. 2017. On the interaction of P-stranding and sluicing in Bulgarian. In *Aspects of Slavic Linguistics: Formal Grammar, Lexicon and Communication*, ed. by O. Mueller-Reichau & M. Guhl, 1–28. Berlin: De Gruyter Mouton.

Ackema, P., & A. Neeleman. 2020. Unifying nominal and verbal inflection: Agreement and feature realization. In *Nominalization: 50 Years on from Chomsky's* Remarks, 29–52. Oxford: OUP. Barros, M., P. Elliot, & G. Thoms. 2014. There is no island repair. https://ling.auf.net/lingbuzz/002100.

- Bošković, Ž. 2006. Case checking versus case assignment and the case of adverbial NPs. *Linguistic Inquiry* 37. 522–533.
- Bošković, Ž. 2008. A minimalist account of genitive of quantification. In *Proc. of FDSL 5 Leipzig*, 270–287.
- Caha, P. 2009. The nanosyntax of case. University of Tromsø dissertation.
- Chomsky, N. 1986. Knowledge of language: Its nature, origin, and use. Westport, CT: Praeger.
- Di Sciullo, A., & E. Williams. 1987. On the definition of word. Cambridge, MA: MIT Press.
- Franks, S. 1995. Parameters of Slavic morphosyntax. New York: OUP.
- Grabovac, A. 2022. *Maximizing the concord domain: Concord as spellout in Slavic*. University College London dissertation.
- Kim, J. 2012. What sluicing comes from in Korean is pseudo-cleft. *Korean Journal of Linguistics* 37. 69–106.
- Kim, J.-B. 2015. Syntactic and semantic identity in Korean sluicing: A direct interpretation approach. *Lingua* 166. 260–293.
- Lasnik, H. 2005. The syntax of silence. By Jason Merchant. Language 81. 259–265.
- McFadden, T. 2004. The position of morphological case in the derivation: A study on the syntax-morphology interface. University of Pennsylvania dissertation.
- Merchant, J. 2001. The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford: OUP.
- Norris, M. 2014. A Theory of Nominal Concord. UC Santa Cruz dissertation.
- Nykiel, J., J.-B. Kim, & R. Sim. 2023. Case-matching effects under clausal ellipsis and the cuebased theory of sentence processing. *J. Linguistics* 59, 327–360.
- Öztürk, B. 2002. Turkish as a non-pro-drop language. In *The verb in Turkish*, ed. by E. Erguvanlı Taylan, 239–259. Amsterdam: John Benjamins.
- Park, M.-K. 2014. The syntax of 'sluicing'/'fragmenting' in Korean: Evidence from the copula -i-'be'. *Linguistic Research* 31. 103–133.
- Pleshak, P. 2023. Brazilian Portuguese exceptives questionnaire. University of Maryland.
- Reeve, M. 2012. Clefts and their relatives. John Benjamins.
- Ross, J. 1969. Guess who? In Papers from the fifth regional meeting of the Chicago Linguistic Society, 252–286.
- Saab, A. 2010. Silent interactions: Spanish TP-ellipsis and the theory of island repair. *Probus* 22. 73–116.
- Stjepanović, S. 2012. Two cases of violation repair under sluicing. In *Sluicing: Cross-linguistic perspectives*, ed. by J. Merchant & A. Simpson, 68–82. Oxford: OUP.
- Stjepanović, S. 2013. Genitive of quantification, ellipsis, and the structure of inherently case-marked noun phrases. In *Proc. of FASL 21 Indiana*, 317–330.
- Thoms, G. 2019. Reassessing case-matching. Handout from Sluicing@50 Workshop. University of Chicago.
- van Craenenbroeck, J. 2017. Sluicing as sharing. In *A schrift to fest Kyle Johnson*, ed. by N. LaCara, K. Moulton, & A.-M. Tessier, 398–399. Amherst, MA: Linguistics Open Access Publications.
- Vicente, L. 2015. Morphological case mismatches under sluicing. *Snippets* 29. 15–17.
- Šarić, A. 2014. Numeral induced agreement mismatches in Serbo-Croatian. Master's thesis, Utrecht University.
- Wechsler, S., & L. Zlatic. 2003. The many faces of agreement. Stanford, CA: CSLI Publications.