J. Linguistics **59** (2023), 229–234. © The Author(s), 2023.
Published by Cambridge University Press
doi:10.1017/S0022226722000585

Introduction to the special issue Experimental and Corpus-based Approaches to Ellipsis

GABRIELA BÎLBÎIE
University of Bucharest & Laboratoire de Linguistique Formelle
JOANNA NYKIEL
Stockholm University

(Received 30 November 2022; revised 20 December 2022)

Ellipsis has been, and continues to be, of both theoretical and empirical interest. It affects the syntax of phrases or clauses by stranding their various constituents but keeps the semantics of the stranded constituents identical to that of their non-elliptical counterparts. The theoretical value of ellipsis lies, therefore, in the relationship between meaning and form that it encodes, such that a complete propositional meaning is paired with what appears to be a syntactically incomplete form. This property of ellipsis has inspired researchers to probe, in particular, the syntax of ellipsis and the role the surrounding context plays in helping resolve ellipsis, as stranded constituents depend on the surrounding context for their interpretation. Among the constructions that have attracted considerable attention over the years are clausal ellipsis (e.g. sluicing (Example 1), sprouting (Example 2), stripping (Example 3), and fragments (Example 4)), pseudogapping (Example 5), gapping (Example 6), and Right Node Raising (RNR) (Example 7), all of which are discussed in the contributions to this special issue.

(1) "My God", he said. "They survived. **How the hell?**" (see Kim & Kim)
(2) Although it's unclear **with what money**, Jill bought a ferry. (see Jung & Goodall)
(3) Mari saw Kadi, **not Anna.** (see Kaps)
(4) A: Whose car got stolen?
B: **Harvey's** (car). (see Nykiel, Kim & Sim)
(5) Ask Doll, who spoke as much about his schoolboy career ending **as he did of the season in general.** (see Poppels & Miller)
(6) Alfonso stole the emeralds, **and Mugsy the pearls.** (see Bîlbîie, de la Fuente & Abeillé)
(7) **Who is** and who should be making the criminal law here? (see Abeillé, Shiraishi & Hemforth)

1 Two main strands of research on ellipsis

One long-standing strand of research on ellipsis centers around the debate about whether a seemingly incomplete form is just the phrase that we see or whether it has more complex syntax at some level of derivation. Proposals that defend the view that ellipsis has no unpronounced structure (referred to as DIRECT INTERPRETATION or NONSTRUCTURAL in the literature) are typically couched in constraint-based frameworks like Head-driven Phrase Structure Grammar (Ginzburg & Sag 2000, Sag & Nykiel 2011, Abeillé et al. 2014, Kim 2015, Bîlbîie 2017, Kim & Abeillé 2019, Abeillé & Kim 2022, Nykiel & Kim 2022a), Simpler Syntax (Culicover and Jackendoff 2005), Construction Grammar (Goldberg & Perek 2018), or Categorial Grammar (Kubota & Levine 2017). Proposals that defend the view that unpronounced sentential structure underlies ellipsis (termed STRUCTURAL in the literature) are couched in various versions of Transformational Grammar (from Ross 1967 and Hankamer & Sag 1976 to Merchant 2001, 2004, 2013, Johnson 2009, van Craenenbroeck 2010, and Weir 2014). The contributions by Jung & Goodall, Kim & Kim, Poppels & Miller, and Nykiel et al. address this debate directly or indirectly (more detail about how ellipsis is treated within these frameworks can be found in van Craenenbroeck & Temmerman 2018 and Nykiel & Kim 2021).

Another popular strand of research focuses on whether the surrounding context determines the syntax of ellipsis and, if so, how. Of particular interest, here, is the degree of structural match that must hold between an ellipsis and its antecedent. Structural accounts of ellipsis require various degrees of STRUCTURAL MATCH (e.g. Merchant 2001, 2013, Chung 2006, 2013), having abandoned a full structural identity requirement once assumed in proposals like Hankamer and Sag (1976) or Sag (1976). Direct interpretation accounts, on the other hand, assume either limited or no structural match between an ellipsis and its antecedent (Ginzburg & Sag 2000, Culicover & Jackendoff 2005, Nykiel & Kim 2022a). The latter accounts typically garner empirical support from evidence that STRUCTURAL MISMATCH is observed in actual usage and judged acceptable in psycholinguistic experimentation. The acceptability of structural mismatch under different types of ellipsis features prominently in the contributions by Poppels & Miller, Abeillé et al., Bîlbîie et al., and Nykiel et al.

2 Advantages of empirical investigations

While the theoretical work on ellipsis has led to important discoveries and avenues for future research, it has also raised questions that are difficult to answer without a solid foundation of empirical data. These questions relate, for instance, to the extent to which various elliptical constructions show island sensitivity (for experimental work, see, e.g. Yoshida et al. 2013, 2014, 2019), connectivity effects (for experimental and corpus-based work, see, e.g. Nykiel 2013, Molimpakis 2019, Nykiel & Hawkins 2020, Lemke 2021, Nykiel & Kim 2022b), or in other ways pattern the way the existing theoretical proposals would lead us to expect. Recent work on

https://doi.org/10.1017/S0022226722000585 Published online by Cambridge University Press

ellipsis has, therefore, seen a marked shift away from reliance on native speaker judgments and toward more controlled corpus-based and experimental methodologies, a shift that is reflected in all of our contributions. The significance of this methodological shift lies specifically in advancing our understanding of the empirical coverage that the theoretical proposals currently have and in highlighting potential challenges some or all of these proposals face when unexpected data need to be accommodated.

Beyond their ability to test the empirical coverage of the theoretical proposals, certain experimental paradigms borrowed from psycholinguistics (e.g. self-paced reading or eye tracking) can offer insight into the online profile of various types of ellipsis. Such paradigms have been recruited in research on ellipsis in the hope that they may help us answer more complex questions surrounding the representation of elliptical structures, including, for instance, the questions of what linguistic information is accessed during the processing of ellipsis and whether accessing this information constitutes evidence for the presence/absence of unpronounced structure at the ellipsis site (see, e.g. Kaan et al. 2004, Yoshida et al. 2012). But even regardless of what questions it is intended to probe, psycholinguistic experimentation helps establish stronger links between theoretical linguistics and psycholinguistics, given the latter's equally long-standing interest in the online processing of the unusual mapping between form and meaning that ellipsis represents (from Murphy 1985 and Tanenhaus & Carlson 1990 to Martin & McElree 2011, Frazier 2013, and Parker 2018). We will see two of the contributions (Kaps and Nykiel et al.) engage, directly or indirectly, with the processing dimension of ellipsis.

3 Contributions to this collection

Below, we explain in more detail how the individual contributions fit into the various strands of ellipsis research outlined above. We have organized the contributions into three sections.

3.1 *The syntax of ellipsis*

Jung and Goodall examine experimentally to what extent backward sprouting (e.g. *Although it's unclear with what money, Jill bought a ferry*.) resembles filler-gap dependencies in terms of sensitivity to clause boundaries (distance) and sensitivity to islands. They find that there is a crucial difference between them: backward sprouting is not sensitive to islands, while filler-gap dependencies are (and both show distance effects). This finding speaks in favor of direct interpretation accounts of ellipsis and has implications for our understanding of island sensitivity.

Kim and Kim's corpus investigation of the aggressively non-D-linked construction (e.g. *Why the hell*?) leads them to propose a direct interpretation account of it, incorporating insights from Ginzburg and Sag's (2000) Head-Driven Phrase Structure Grammar (HPSG) account of clausal ellipsis, that is, fragments and sluices. Kim

https://doi.org/10.1017/S0022226722000585 Published online by Cambridge University Press

& Kim's data come from the Corpus of Contemporary American English and include the aggressively non-D-linked construction in its uses with and without ellipsis.

Poppels and Miller's contribution is an experimental investigation of connectivity effects under pseudogapping. They examine the acceptability of PP remnants with matching and mismatching correlates in the antecedent clause (e.g. *Ask Doll, who spoke as much about his schoolboy career ending as he did of the season in general.*). They find that utterances with mismatching PPs across the ellipsis clause and its antecedent are degraded relative to their preposition-matched counterparts, and that the same effect arises in non-elliptical controls. Upon examining the cause of this effect, they conclude that the data are more naturally accounted for under a direct interpretation approach to pseudogapping.

Nykiel, Kim and Sim investigate one kind of connectivity effect, case-matching effects, under clausal ellipsis (e.g. A: *Whose car got stolen*? B: *Harvey's (car).*). They offer novel experimental evidence for a case-matching preference in Korean when a fragment and its correlate may differ in case marking. This case-matching preference corresponds to a known case of mandatory case matching in Hungarian, but their relationship is not predicted by any of the existing direct interpretation or structural accounts of case-matching effects under clausal ellipsis. Nykiel, Kim & Sim's proposal is to derive mandatory and optional case matching from the predictions of the cue-based theory of sentence processing.

3.2 *The role of the surrounding context*

Abeillé, Shiraishi and Hemforth's contribution investigates voice mismatch in Right-Node Raising (e.g. *Please tell me who has and who was shaved.*), based on French data. It does so in two stages: first a corpus study (frTenTen 2012) is employed to characterize the distribution of RNR with voice mismatch, and then two acceptability rating experiments explore how structural mismatch and semantic contrast affect acceptability. Abeillé, Shiraishi and Hemforth find no penalty for voice mismatch in French but a penalty for the absence of contrast between the two conjoined clauses. An HPSG deletion-based analysis is presented that requires lexeme identity for RNR to be acceptable.

Bîlbîie, de la Fuente and Abeillé examine experimentally Johnson's "No Embedding Constraint" proposed for gapping in English (e.g. **Alfonso stole the emeralds, and I think (that) Mugsy the pearls*.). A series of three acceptability experiments, designed to measure the interaction between complementizer omission and factivity in English embedded complement clauses, shows that the acceptability of embedded gapping is variable, depending on two factors: (i) the absence of the complementizer *that* increases acceptability, and (ii) acceptability increases depending on whether the governing predicate is non-factive, semifactive versus true factive. These new data lead the authors to argue for a constructionist fragment-based analysis, where the gapped clause is a non-finite phrase that has to address the same Question-Under-Discussion (QUD) as its source clause. The

https://doi.org/10.1017/S0022226722000585 Published online by Cambridge University Press

authors conclude that English has the same sensitivity to the semantic class of the governing predicate as other languages but that the requirements on the presence/absence of the complementizer *that* are English-specific.

3.3 *Processing of ellipsis*

Kaps offers novel experimental evidence for information-structural effects that underlie a well-documented locality preference under clausal ellipsis with contrastive remnants (e.g. *Mari saw Kadi, not Anna.*). Her evidence comes from Estonian, where flexible word order makes it possible to tease apart information-structural effects from locality effects. This property of Estonian offers a new way of testing whether focus-marked correlates are easier to retrieve regardless of the distance from the ellipsis site, as predicted by Harris & Carlson's (2018) Information Structure hypothesis and confirmed by Kaps in an eye-tracking experiment.

To conclude, this collection of papers brings experimental and corpus-based results to bear on topics within the domain of ellipsis where judgments are subtle and where traditional informal methods of collecting data have underestimated the space of acceptability. We hope that this special issue will inspire further research on the elliptical phenomena discussed here and help extend experimental and quantitative methods to other areas of ellipsis.

REFERENCES

Abeillé, Anne, Gabriela Bîlbîie & François Mouret. 2014. A Romance perspective on gapping constructions. In Hans Boas & Francesco Gonzálvez-García (eds.), *Romance perspectives on Construction Grammar*, 227–267. Amsterdam/Philadelphia: John Benjamins.

Abeillé, Anne & Jong-Bok Kim. 2022. 'Me too' fragments in English and French: A direct interpretation approach. *The Linguistic Review* 39.3, 495–524.

Bîlbîie, Gabriela. 2017. *Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français*. Berlin: Language Science Press.

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* 31, 73–91. Berkeley, CA.

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

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

Frazier, Lyn. 2013. A recycling approach to processing ellipsis. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosis in syntax*, 485–501. Oxford: Oxford University Press.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations. The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.

Goldberg, Adele & Florian Perek. 2018. Ellipsis in Construction Grammar. In Jeroen van Craenenbroeck & Tanja Temmerman (eds.), *The Oxford Handbook of Ellipsis*, 188–204. Oxford University Press.

Hankamer, Jorge & Ivan A. Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7, 391–428.

Harris, Jesse A. & Katy Carlson. 2018. Information structure preferences in focus-sensitive ellipsis: How defaults persist. *Language and Speech* 61.3, 480–512.

Johnson, Kyle. 2009. Gapping is not (VP)-ellipsis. *Linguistic Inquiry* 40, 289–328.

Kaan, Edith, Frank Wijnen & Tamara Y. Swaab. 2004. Gapping: Electrophysiological evidence for immediate processing of 'missing' verbs in sentence comprehension. *Brain and Language* 89, 584–592.

Kim, Jong-Bok. 2015. Syntactic and semantic identity in Korean sluicing: A direct interpretation approach. *Lingua* 166, 260–293.

Kim, Jong-Bok & Anne Abeillé. 2019. *Why*-stripping in English. *Linguistic Research* 36, 365–387.
Kubota, Yusuke & Robert Levine. 2017. Pseudogapping as pseudo-VP ellipsis. *Linguistic Inquiry* 48.2, 212–257.
Lemke, Robin. 2021. *Experimental investigations on the syntax and usage of fragments*. Berlin: Language Science Press.
Martin, Andrea E. & Brian McElree. 2011. Direct-access retrieval during sentence comprehension: Evidence from sluicing. *Journal of Memory and Language* 64, 327–343.
Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27, 661–738.
Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44, 77–108.
Molimpakis, Emilia. 2019. *Accepting preposition-stranding under sluicing cross-linguistically: A noisy-channel approach.* Ph.D. dissertation, University College London.
Murphy, Gregory L. 1985. Psychological explanations of deep and surface anaphora. *Journal of Pragmatics* 9, 785–813.
Nykiel, Joanna. 2013. Clefts and preposition omission in sluicing. *Lingua* 123, 74–117.
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. 2021. Ellipsis. In Stefan Müller, Anne Abeillé, Robert D. Borsley & Jean-Pierre Koenig (eds.), *Head-driven Phrase Structure Grammar: The handbook*, 843–883. Berlin: Language Science Press.
Nykiel, Joanna & Jong-Bok Kim. 2022a. Fragments and structural identity on a direct interpretation approach. *Journal of Linguistics* 58.1, 73–109.
Nykiel, Joanna & Jong-Bok Kim. 2022b. On the grammaticality of morphosyntactically reduced remnants in Polish sluicing. *Linguistics* 60.1, 177–213.
Parker, Dan. 2018. A memory-based explanation of antecedent-ellipsis mismatches: New insights from computational modeling. *Glossa* 3.129, 1–27.
Ross, John Robert. 1967. *Constraints on variables in syntax*. dissertation, MIT.
Sag, Ivan A. 1976. *Deletion and logical form*. Ph.D. dissertation, MIT.
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.
Tanenhaus, Michael K. & Greg N. Carlson. 1990. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes* 5, 257–280.
van Craenenbroeck, Jeroen. 2010. *The syntax of ellipsis*. Oxford: Oxford University Press.
van Craenenbroeck, Jeroen & Tanja Temmerman. 2018. *The Oxford Handbook of Ellipsis*. Oxford University Press.
Weir, Andrew. 2014. *Fragments and clausal ellipsis.* Ph.D. dissertation, University of Massachusetts Amherst.
Yoshida, Masaya, Michael Walsh Dickey & Patrick Sturt. 2012. Predictive processing of syntactic structure: Sluicing and ellipsis in real-time sentence processing. *Language and Cognitive Processes* 28.3, 272–302
Yoshida, Masaya, Jiyeon Lee & Michael Walsh Dickey. 2013. The island (in)sensitivity of sluicing and sprouting. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental syntax and island effects*, 360–376. Cambridge: Cambridge University Press.
Yoshida, Masaya, Lauren Ackerman, Rebekah Ward & Morgan Purrier. 2014. The processing of backward sluicing and island constraint. *Proceedings of the 43rd Annual Meeting of the North East Linguistic Society (NELS* 43*)*, 261–272.
Yoshida, Masaya, David Potter & Tim Hunter. 2019. Condition C reconstruction, clausal ellipsis and island repair. *Natural Language and Linguistic Theory* 37, 1515–1544.

Authors' addresses: *(Bîlbîie)*
University of Bucharest & Laboratoire de Linguistique Formelle

(Nykiel)
Stockholm University

https://doi.org/10.1017/S0022226722000585 Published online by Cambridge University Press