



Katharina Ehret, Aleksandrs Berdicevskis*, Christian Bentz and Alice Blumenthal-Dramé

Measuring language complexity: challenges and opportunities

<https://doi.org/10.1515/lingvan-2022-0133>

Received November 7, 2022; accepted March 9, 2023; published online May 4, 2023

Abstract: This special issue focuses on measuring language complexity. The contributions address methodological challenges, discuss implications for theoretical research, and use complexity measurements for testing theoretical claims. In this introductory article, we explain what knowledge can be gained from quantifying complexity. We then describe a workshop and a shared task which were our attempt to develop a systematic approach to the challenge of finding appropriate and valid measures, and which inspired this special issue. We summarize the contributions focusing on the findings which can be related to the most prominent debates in linguistic complexity research.

Keywords: complexity metrics; computational linguistics; corpus linguistics; language complexity; typology

1 Introduction

For the past two decades, language complexity has been a widely discussed – and sometimes hotly debated – topic, engaging researchers from diverse linguistic disciplines such as language typology, language evolution, second language acquisition (SLA), computational linguistics, and psycho- and neurolinguistics (e.g. Armeni et al. 2017; Baechler and Guido 2016; Baerman et al. 2015; Futrell and Hahn 2022; Housen et al. 2019; Kortmann and Szmrecsanyi 2012; Mufwene et al. 2017). In theoretical complexity research, major questions under debate include, among others, whether all languages are equally complex, the extent to which language complexity is influenced by socio-demographic variables such as population size or the number of adult learners, and the extent to which complexity correlates with processing difficulty.

Any theoretical questions about complexity necessarily generate a need to measure complexity. Naturally, measuring language complexity involves numerous methodological and theoretical challenges such as, for instance, (a) the applicability of complexity measures to different data types (e.g. descriptive grammars, typological databases, corpora), (b) the extent to which different complexity measures correlate, (c) how and whether complexity metrics can be fruitfully applied to different levels of language description (e.g. morphology and syntax), and (d) how operationalizations of language complexity relate to underlying constructs (i.e. what does a given metric actually measure?).

In search for answers to these questions, we convened the “Interactive Workshop on Measuring Language Complexity” in September 2019 and invited participants to take part in a shared task on measuring language complexity (see Section 3). In the same spirit, the current special issue brings together empirical and theoretical contributions from the perspectives of theoretical typology, computational linguistics, and SLA research which address some of the outlined challenges.

This introductory article is structured as follows. In Section 2, we provide a brief review of common approaches to measuring complexity and discuss some of the challenges related to this endeavour. Section 3 describes the shared task on which some of the contributions in this special issue are based. In Section 4, the individual contributions are presented and summarized. Section 5 concludes by highlighting some of the findings presented in this issue and discussing them in light of prominent debates in the linguistic complexity community.

*Corresponding author: Aleksandrs Berdicevskis, Språkbanken Text, Department of Swedish, Multilingualism, Language Technology, University of Gothenburg, Gothenburg, Sweden, E-mail: aleksandrs.berdicevskis@gu.se

Katharina Ehret and Alice Blumenthal-Dramé, Department of English, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany

Christian Bentz, Department of General Linguistics, University of Tübingen, Tübingen, Germany

2 Measuring language complexity

Measures of language complexity exist in abundance. In the theoretical complexity literature, they can broadly be distinguished either as measuring *absolute* complexity or *relative* complexity (Miestamo 2008, see also Bulté and Housen 2012). Absolute complexity measures address language complexity inherent in a linguistic system and are variously operationalized as, for instance, the number of rules in a grammar (McWhorter 2012), the number of irregular markers in a linguistic system (Trudgill 2011), or, in information-theoretic terms, as the shortest possible description of a naturalistic text sample (Ehret 2021; Juola 2008). In contrast, relative complexity measures address language complexity in relation to language users, and are often equated with SLA difficulty or, more generally, with “cost and difficulty” (Dahl 2004). Relative complexity measures can be operationalized as, for example, the number of markers which are difficult to acquire for second language learners (Kusters 2008), or in terms of processing efficiency (Hawkins 2009). That said, the distinction between absolute and relative measures is frequently more fuzzy than clear-cut. Often there is a substantial overlap, and sometimes there is a mismatch between what is being explored and what is being measured (for a detailed discussion, see Ehret et al. 2021).

A second major distinction is generally made between *local* and *overall/global* measures (Kortmann and Schröter 2020; Miestamo 2008). The vast majority of existing complexity measures address language complexity at local linguistic levels, such as morphology, syntax, or phonology. Some, mainly information-theoretic, measures have also been proposed to address language complexity at an overall structural level.

In SLA research, the notion of language complexity also plays an important role. Yet, like in theoretical complexity research, there is disagreement when it comes to defining complexity, and calls have been made for precise and clear-cut definitions (for details, see e.g. Bulté and Housen 2012; Ortega 2012). The motivation for measuring language complexity in SLA research is threefold: first, “to gauge proficiency”, second, “to describe performance”, and third, “to benchmark development” (Ortega 2012:128). Thus, most measures of language complexity target complexity at the syntactic or lexical level, and are variously operationalized as measures of unit length (e.g. T-unit length), subordination frequency, or measures of the frequency of “complex” forms (see Ortega 2003). Only recently have measures addressing other linguistic levels (e.g. morphology) been proposed (Brezina and Pallotti 2019; for details, see also Kuiken 2022, this issue).

Be that as it may, despite the panoply of measures in existence, neither theoretical linguists nor SLA researchers agree on a common measure of complexity. This poses several methodological and theoretical challenges.

First, the extent to which different studies – for example, complexity rankings of languages – can be compared. Drawing conclusions or making generalizations across different studies is only fruitful and feasible if the complexity measures can be compared. One question related to this challenge is whether measures that are supposed to gauge the same type of complexity yield similar results. Recent data-driven studies report that morphological complexity measures correlate significantly (Çöltekin and Rama 2022 this volume, Berdicevskis et al. 2018), while there seems to be less agreement between syntactic measures (Berdicevskis et al. 2018). Another question is how to establish that differences observed between various languages, corpora, or measures are large enough to be of interest. Ehret et al. (2021), for instance, suggest to steer away from pure significance testing, and rather establish effect sizes for shifts in complexity distributions.

Second, the extent to which complexity measures capture the underlying theoretical constructs they are supposed to capture is often unclear (Bulté and Housen 2012; Norris and Ortega 2009:26–28). For example, some measures operationalize more than one linguistic level by, for example, mixing syntax and morphology (Bulté and Housen 2012:29); while some measures which are meant to address several levels, merely operationalize one (Norris and Ortega 2009:560). In both cases, the measures do not actually reflect the underlying theoretical construct. This question, so far, has primarily been discussed in the SLA literature (Bulté and Housen 2012), yet, it is also often unclear whether complexity measures in theoretical linguistics actually operationalize what they advertise. Consider, for instance, a measure of syntactic complexity which operationalizes “syntax” as “degree of subordination”. While such a measure can undoubtedly be interesting and useful, it is questionable whether it could be used as an operationalization of overall syntactic complexity, since it arguably does not capture the construct “syntax” as a whole.

Third, how reliable are the numerical values given by a measure? The core problem is that we do not actually know the complexity of individual languages (or certain parts of linguistic systems). In the absence of a ground truth (gold standard), it is extremely difficult to evaluate the measures. Berdicevskis et al. (2018) propose two indirect ways to evaluate corpus-based measures in the absence of a gold standard. They gauge the *robustness* of measures (how much the values change when the same measure is applied to different samples drawn from the same corpus) and their *validity* (whether measures yield more similar values when applied to very similar languages than when applied to languages known to be different).

Of course, this special issue cannot solve all of these challenges. The authors and the editors, however, tried to keep them in mind while measuring complexity, and some articles actually address the outlined questions explicitly.

3 The shared task

Four of the contributions in this special issue are based on a shared task which was featured at the “Interactive Workshop on Measuring Language Complexity” at the Freiburg Institute for Advanced Studies in September 2019. The basic idea of shared tasks is to address particularly challenging problems which are hard for individual researchers to tackle. Shared tasks are very common in natural language processing, much less so in linguistics per se. The main reason is perhaps that for theoretical questions, the ground truth is usually unknown, and it is thus impossible to construct a benchmark against which the results could be compared. This is true also for measuring complexity. Nonetheless, the shared task provided an empirically and methodologically well-grounded basis for addressing theoretical and methodological challenges of measuring language complexity as outlined in Section 2. The idea of the shared task was to compare different measures of language complexity when they are applied to the same data sets, and look for solutions for the outlined challenges.

In this vein, participants applied their own measure(s) of language complexity to either a parallel text database (for raw-text-based metrics), or a non-parallel annotated text database (for text-and-annotation-based metrics). The parallel text database comprised a subsample¹ of the Parallel Bible Corpus (Mayer and Cysouw 2014), covering 49 typologically diverse languages which are part of the 100-language sample in the *World Atlas of Language Structures* (Dryer and Haspelmath 2013). The annotated non-parallel database, on the other hand, covered 44 treebanks from the Universal Dependencies (UD) corpora version 2.3 (Nivre et al. 2018). The results² originating from the shared task are thus comparable, at least in the sense of being based on the same data set(s).

In total, 31 different measures and some variants thereof were applied. They target language complexity at various linguistic levels – morphology, syntax, and the lexicon – or assess language complexity in terms of information density. Despite the fact that these measures operationalize the various linguistic levels differently, for example, by measuring different units of analysis (such as words vs. *n*-grams), it turned out that often measures yield similar, that is highly correlated, results (see Bentz et al. 2016 and Çöltekin and Rama 2022 in this special issue). We take this as an indicator for similar conceptualizations of language complexity at various linguistic levels.

4 Overview of the special issue

In the following, we present an overview and summaries of the various contributions to this special issue, distinguishing between contributions which are entirely or partially based on the shared task data set(s) (Section 4.1) and other empirical and theoretical contributions on measuring language complexity (Section 4.2).³ The summaries are otherwise presented in alphabetical order.

¹ Available at http://www.christianbentz.de/MLC2019_data.html.

² Available at <https://github.com/IWMLC/language-complexity-metrics>.

³ Most of the contributions based on other data sets were submitted via a separate call for papers. For those, using the shared task data set was optional as they were not presented at the workshop.

4.1 Contributions based on the shared task

Bentz, Gutierrez-Vasques, Sozinova, and Samardžić offer a corpus-based meta-analysis of 80 typologically diverse languages across 28 different morphosyntactic complexity measures used in the shared task. To begin with, they assess trade-offs (i.e. negative correlations) and agreement (i.e. positive correlations) between different measures. Within the same domain (i.e. morphology or syntax) measures tend to positively correlate, while between domains there is a tendency for negative correlations. Interestingly, this result was in accordance with previous expectations; see, for example, Sinnemäki (2011:60). Another finding of the paper contributes to the debated equi-complexity hypothesis which claims that, overall, all languages are equally complex (Hockett 1958). Analysing a subsample of 44 languages, Bentz et al. observe no statistically significant location shifts in a pairwise comparison of complexity vectors. In other words, the overall morphosyntactic complexity of the analysed languages seems virtually identical. Thus, the null hypothesis of equi-complexity cannot be rejected.

Brunato and Venturi investigate structural complexity at the syntactic level in 63 UD treebanks. Using linguistic profiling, they analyse the distribution of syntactic features and their complexity. A hierarchical cluster analysis shows that the analysed languages mostly cluster according to the expected typological classification. Yet, some outliers illustrate that register affects syntactic complexity.

The contribution by **Çöltekin and Rama** evaluates the degree of similarity between eight corpus-based measures of morphological complexity; some of them, such as type-token ratio and mean size of paradigm, are rather well studied, some, such as inflection accuracy, are rather novel. Using correlation and principal component analyses, Çöltekin and Rama find that the eight measures do not only yield intuitive rankings of the analysed languages but also correlate to a large degree. Their principal component analysis suggests that all eight measures assess the same underlying construct.

Sinnemäki and Haakana present two case studies on morphological complexity in possessive noun phrases. In the first case study they analyse the relationship between head marking and dependent marking in a typological database of 316 languages: their results show that there is an inverse relationship between head marking and dependent marking both within and across languages. In the second case study, a corpus analysis of 44 UD treebanks, they test whether presence of overt morphological marking (morphological complexity) affects dependency length (syntactic complexity), but find no robust cross-linguistic trend.

4.2 Other empirical and theoretical contributions

Grieve proposes a new perspective on analysing complexity: based on Biberian register analysis, he suggests evaluating a language's complexity in terms of situational diversity. In short, languages which are used in a broader range of communicative contexts are considered more situationally complex. Grieve further argues that, over time, increasing situational complexity should lead to decreasing grammatical complexity as simpler grammars would be more adaptable to different communicative contexts.

Kuiken provides an overview of complexity research in SLA, touching upon key issues in the discipline such as the definition of the construct “complexity” and its measurement. He concludes, *inter alia*, with a call for valid and reliable measures as well as developmental measures which can trace performance at all levels of proficiency.

Leufkens showcases how syntagmatic redundancy – which refers to constructions that are represented by more than one form such as negative concord – can be related to metrics of absolute and relative complexity. While redundancy increases absolute complexity, Leufkens claims that redundant marking facilitates both processing and acquisition. Furthermore, in an analysis of concord across 50 distinct languages, Leufkens finds that isolating languages exhibit little to no concord – a finding which supports evidence against redundancy being universal in language.

Liu's contribution presents a mixed-effects logistic regression analysis of 20 typologically diverse treebanks analysing the impact of four well-known constraints on syntactic word order choice. Among the four constraints, dependency length emerges as the strongest predictor: there is a cross-linguistic tendency for shorter constituents

to appear closer to the syntactic head on the phrasal level. The impact of the other predictors is rather weak and the results hint at the fact that factors other than those analysed may play a role (e.g. givenness).

Malaia, Borneman, Kurtoglu, S. Gurbuz, Griffin, Crawford, and A. Gurbuz contribute a sign-language perspective to measuring language complexity. While working with sign languages is still a challenge – they lack a common system of transliteration – modern recording technologies permit the analysis of spatio-temporal parameters of sign signals. The paper reviews the applicability of multiple complexity metrics (such as entropy-based metrics). It illustrates that the modelling of information parameters of sign signals and sign processing exhibits parallels to spoken languages: among others, sign languages can be characterized as complex systems subject to power laws (e.g. Zipf's law). The authors conclude by stressing that the development of sensitive and robust complexity metrics to analyse sign languages are crucial, for instance, for sign recognition and translation.

Morozova, Escher, and Rusakov study absolute complexity in the phonological and morphological inventory of 919 Slavic varieties. Overall, their analysis confirms previous classifications of these varieties, which fall into two large areas: the Serbo-Croatian varieties, which exhibit more complexity; and Bulgarian-Macedonian varieties, which tend towards lower complexity. They furthermore show that complexity correlates with proximity to the Albanian border (presumably due to a complexifying type of contact with Albanian) and with altitude (presumably reflecting the tendency of highland societies to be more isolated and thus more prone to preserve complex features). The authors show how different contact scenarios can result in complexification and maintenance of complex features on the one hand, and in loss and simplification on the other hand.

Shcherbakova, Gast, Blasi, Skirgård, Gray, and Greenhill present a large-scale morphosyntactic analysis of 368 languages and investigate the relation between grammatical information encoded by verbs and nouns using phylogenetic modelling. On the global scale, they find weak positive correlations, while they also observe trade-offs for certain combinations of features. They also find a global trade-off in Indo-European languages, which suggests that accretion and loss of nominal and verbal complexity is lineage-specific. In the light of the equi-complexity hypothesis then, these findings support claims that languages differ in the amount of grammatical information encoded.

5 Conclusions

We believe that some important theoretical and empirical conclusions can be drawn from the contributions to this special issue, the shared task, and the discussions at the “Interactive Workshop on Measuring Language Complexity”.

One of the questions raised at the workshop was how well different complexity measures correlate, or, in other words, whether measures that are supposed to measure the same type of complexity do indeed yield similar results. As exemplified by Çoltekin and Rama's contribution, there is considerable agreement between different operationalizations of complexity at the linguistic level of morphology. This strongly suggests that there is a common understanding of the construct of complexity in this domain which can help to establish similar agreed-upon conceptualizations of complexity at other linguistic levels.

Another important finding to be considered in future complexity research is the relevance of register. Grieve urges taking “situational diversity” into account when measuring language complexity. From this viewpoint, languages which are used in more diverse situations – and thus have a higher register diversity – should be considered more situationally complex. This level of complexity is then expected to negatively correlate with grammatical complexity. The importance of registers also surfaces in Brunato and Venturi's contribution. They find that register variation is systematically linked to their measurements of syntactic complexity. This influence of register on language complexity has been largely overlooked in theoretical complexity research, although it has been previously noted that complexity varies within languages depending on the analysed register (Ehret 2021; Szmrecsanyi 2009).

On a more theoretical plane, the contributions in this special issue touch upon two controversial topics. First, whether language contact, specifically adult SLA, leads to simplification, and second, whether all languages are overall equally complex. The theoretical complexity literature provides controversial evidence for the fact that

adult SLA leads to simplification (Koplenig 2019; Nichols 2009; Trudgill 2011). In light of this controversy, Morozova et al.'s findings are particularly noteworthy as they corroborate the contact scenarios outlined in Trudgill (2011) and show that under certain circumstances, language contact leads to the loss of complex features.

The question whether, overall, all languages are equally complex has always been controversial. In the past, this was certainly related to (d)evaluative judgements about “simple” and “complex” languages. However, we hold that “simple” and “complex” are scientific terms, subject to rigorous definitions and measurements, without any further judgement of value. That said, there are several publications in this special issue which relate to the equi-complexity hypothesis. For instance, Bentz et al. report the absence of significant location shifts comparing the morphosyntactic complexity vectors of 44 languages, and hence claim that the equi-complexity hypothesis cannot be rejected. Shcherbakova et al. find trade-offs between nominal and verbal inflectional complexity in Indo-European languages, but positive correlations between these domains in Sino-Tibetan languages. Thus, trade-offs are here argued to be lineage-specific rather than a general tendency across families. Sinnemäki and Haakana's finding that there is the inverse relationship between head marking and dependent marking could also be interpreted as evidence for a trade-off. In fact, such cross-domain trade-offs, especially between morphology and syntax, are quite common in the complexity literature (e.g. Ehret 2021; Koplenig et al. 2017; Sinnemäki 2011), yet they do not provide straightforward evidence for the overall equal complexity of languages (see Bentz et al. in this issue, and Fenk-Oczlon and Fenk 2014 for an extensive discussion). Nevertheless, such trade-offs are potentially relevant from the perspective of Zipf's principle of least effort (1949): information encoded morphologically does not have to be encoded syntactically – and vice versa. This might also play into the observation that language users keep the amount of information conveyed relatively constant across different spoken languages (see e.g. Coupé et al. 2019).

Finally, we believe that shared tasks as the one described here are underrated in linguistics. Despite the fact that only a handful of the original workshop/shared task contributions appear in this special issue, they demonstrate that shared tasks do have a potential to address methodological and theoretical challenges in the field.

Acknowledgements: We would like to thank the *Linguistics Vanguard* editors Kaius Sinnemäki and Jeff Good, who took care of this special issue; two anonymous reviewers who helped us to improve this introduction; and all the contributors and reviewers who made this issue possible.

Research funding: KE gratefully acknowledges funding by the German Research Foundation (DFG [Deutsche Forschungsgemeinschaft], grant no. EH 532/1-1), for the “Interactive Workshop on Measuring Language Complexity”. AB was supported by the Cassandra project (funded by the Marcus and Amalia Wallenberg Foundation [Marcus och Amalia Wallenbergs minnesfond], donation letter 2020.0060) and Swedish national research infrastructure Nationella språkbanken (funded jointly by the Swedish Research Council [Vetenskapsrådet] [2018–2024, contract 2017–00626] and the 10 participating partner institutions). CB was supported by a Swiss National Science Foundation (Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung) grant (No. 176305) “Non-randomness in Morphological Diversity: A Computational Approach Based on Multilingual Corpora”, as well as by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG FOR 2237: Project “Words, Bones, Genes, Tools: Tracking Linguistic, Cultural, and Biological Trajectories of the Human Past”). Our thanks also go to the Freiburg Institute for Advanced Studies (Freiburg Institute for Advanced Studies, Albert-Ludwigs-Universität Freiburg) for a “Junior Researcher Conferences” grant and for hosting the workshop, and the DFG Center for Advanced Studies “Words, Bones, Genes, Tools” in Tübingen, for supporting the workshop.

References

- Armeni, Kristijan, Roel M. Willems & Stefan L. Frank. 2017. Probabilistic language models in cognitive neuroscience: Promises and pitfalls. *Neuroscience & Biobehavioral Reviews* 83, 579–588.
- Baechler, Raffaella & Guido Seiler (eds.), 2016. *Complexity, isolation, and variation*. Berlin: De Gruyter.

- Baerman, Matthew, Dunston Brown & Greville G. Corbett (eds.), 2015. *Understanding and measuring morphological complexity*. New York: Oxford University Press.
- Bentz, Christian, Tatyana Ruzsics, Alexander Koplenig & Tanja Samardžić. 2016. A comparison between morphological complexity measures: Typological data vs. language corpora. In *Proceedings of the workshop on computational linguistics for linguistic complexity (CL4LC)*. Osaka, Japan. Available at: <http://www.aclweb.org/anthology/W16-4117>.
- Berdicevskis, Aleksandrs, Çağrı Çöltekin, Katharina Ehret, Kilu von Prince, Daniel Ross, Bill Thompson, Chunxiao Yan, Vera Demberg, Gary Lupyan, Taraka Rama & Christian Bentz. 2018. Using universal dependencies in cross-linguistic complexity research. In *Proceedings of the second workshop on universal dependencies (UDW 2018)*, 8–17. Association for Computational Linguistics.
- Brezina, Vaclav & Gabriele Pallotti. 2019. Morphological complexity in written L2 texts. *Second Language Research*, 35(1), 99–119.
- Bulté, Bram & Alex Housen. 2012. Defining and operationalising L2 complexity. In Alex Housen, Folkert Kuiken & Ineke Vedder (eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 21–46. Amsterdam: John Benjamins.
- Çöltekin, Çağrı & Taraka Rama. 2022. What do complexity measures measure? Correlating and validating corpus-based measures of morphological complexity. *Linguistics Vanguard*. <https://doi.org/10.1515/lingvan-2021-0007>. <https://www.degruyter.com/document/doi/10.1515/lingvan-2021-0007/html>.
- Coupé, Christophe, Yoon Mi Oh, Dan Dediu, & François Pellegrino. 2019. Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, 5(9), eaaw2594.
- Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.
- Dryer, Matthew S. & Martin Haspelmath (eds.), 2013. *WALS online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at: <https://wals.info/>.
- Ehret, Katharina. 2021. An information-theoretic view on language complexity and register variation: Compressing naturalistic corpus data. *Corpus Linguistics and Linguistic Theory*, 17(2), 383–410.
- Ehret, Katharina, Alice Blumenthal-Dramé, Christian Bentz & Aleksandrs Berdicevskis. 2021. Meaning and measures: Interpreting and evaluating complexity metrics. *Frontiers in Communication*, 6, 640510.
- Fenk-Oczlon, Gertraud & August Fenk. 2014. Complexity trade-offs do not prove the equal complexity hypothesis. *Poznań Studies in Contemporary Linguistics*, 50(2), 145–155.
- Futrell, Richard & Michael Hahn. 2022. Information theory as a bridge between language function and language form. *Frontiers in Communication* 7. <https://doi.org/10.3389/fcomm.2022.657725>. <https://www.frontiersin.org/articles/10.3389/fcomm.2022.657725/full>.
- Hawkins, John A. 2009. An efficiency theory of complexity and related phenomena. In Geoffrey Sampson, David Gil & Trudgill Peter (eds.), *Language complexity as an evolving variable*, 252–268. Oxford: Oxford University Press.
- Hockett, Charles Francis. 1958. *A course in modern linguistics*. New York: Macmillan.
- Housen, Alex, Bastien De Clercq, Folkert Kuiken & Ineke Vedde. 2019. Multiple approaches to complexity in second language research. *Second Language Research*, 35(1), 3–21.
- Juola, Patrick. 2008. Assessing linguistic complexity. In Fred Karlsson, Matti Miestamo & Kaius Sinnemäki (eds.), *Language complexity: Typology, contact, change*, 89–107. Amsterdam: John Benjamins.
- Koplenig, Alexander. 2019. Language structure is influenced by the number of speakers but seemingly not by the proportion of non-native speakers. *Royal Society Open Science*, 6(2), 181274.
- Koplenig, Alexander, Peter Meyer, Sascha Wolfer & Carolin Müller-Spitzer. 2017. The statistical trade-off between word order and word structure—large-scale evidence for the principle of least effort. *PLoS One*, 12(3), e0173614.
- Kortmann, Bernd & Verena Schröter. 2020. Linguistic complexity. In Mark Aranoff (ed.), *Oxford bibliographies in linguistics*. Oxford: Oxford University Press.
- Kortmann, Bernd & Benedikt Szmrecsanyi (eds.), 2012. *Linguistic complexity: Second language acquisition, indigenization, contact (Lingua & Litterae)*. Berlin: Walter de Gruyter.
- Kuiken, Folkert. 2022. Linguistic complexity in second language acquisition. *Linguistics Vanguard*. <https://doi.org/https://doi.org/10.1515/lingvan-2021-0112>.
- Kusters, Wouter. 2008. Complexity in linguistic theory, language learning and language change. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 3–21. Amsterdam: John Benjamins.
- Mayer, Thomas & Michael Cysouw. 2014. Creating a massively parallel Bible corpus. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 3158–3163. Reykjavik, Iceland: European Language Resources Association (ELRA). Available at: http://www.lrec-conf.org/proceedings/lrec2014/pdf/220_Paper.pdf.
- McWhorter, John. 2012. Complexity hotspot: The copula in Saramaccan and its implications. In Bernd Kortmann & Benedikt Szmrecsanyi (eds.), *Linguistic complexity: Second language acquisition, indigenization, contact (Linguae & Litterae)*, 243–246. Berlin: Walter de Gruyter.
- Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 23–41. Amsterdam: John Benjamins.
- Mufwene, Salikoko, Christophe Coupé & François Pellegrino. 2017. *Complexity in language: Developmental and evolutionary perspectives*. Cambridge & New York: Cambridge University Press.
- Nichols, Johanna. 2009. Linguistic complexity: A comprehensive definition and survey. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 64–79. Oxford: Oxford University Press.
- Nivre, Joakim, Mitchell Abrams, Željko Agić, Lars Ahrenberg, Lene Antonsen, Katya Aplonova, Maria Jesus Aranzabe, Gashaw Arutie, Masayuki Asahara, Luma Ateyah, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Elena Badmaeva, Miguel Ballesteros,

- Esha Banerjee, Sebastian Bank, Verginica Mititelu Barbu, Victoria Basmov, John Bauer, et al. 2018. *Universal dependencies 2.3*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. Available at: <http://hdl.handle.net/11234/1-2895>.
- Norris, John M. & Lourdes Ortega. 2009. Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578.
- Ortega, Lourdes. 2003. Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492–518.
- Ortega, Lourdes. 2012. Interlanguage complexity. In Bernd Kortmann & Benedikt Szmrecsanyi (eds.), *Linguistic complexity: Second language acquisition, indigenization, contact (Lingua & Litterae)*, 127–155. Berlin: Walter de Gruyter.
- Sinnemäki, Kaius. 2011. *Language universals and linguistic complexity: Three case studies in core argument marking*. Helsinki: University of Helsinki PhD dissertation.
- Szmrecsanyi, Benedikt. 2009. Typological parameters of intralingual variability: Grammatical analyticity versus syntheticity in varieties of English. *Language Variation and Change*, 21(3), 319–353.
- Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford; New York: Oxford University Press.