



Complexity and Its Relation to Variation

Diana Forker*

Institute of Slavonic Languages and Caucasus Studies, University of Jena, Jena, Germany

This paper is concerned with the relationship between complexity and variation. The main goal is to lay out the conceptual foundations and to develop and systematize reasonable hypotheses such as to set out concrete research questions for future investigations. I first compare how complexity and variation have synchronically been studied and what kinds of questions have been asked in those studies. Departing from earlier surveys of different definitions of complexity, here I classify the majority of complexity studies into two broad types based on two ways of defining this concept. The first type determines and measures linguistic complexity by counting numbers of items (e.g., linguistic forms or rules and interactions between forms). The second type makes use of transparency and the principle of One-Meaning–One-Form. In addition, linguistic complexity has been defined by means of concepts from information theory, namely in terms of description length or information content, but those studies are in the minority. Then I define linguistic variation as a situation when two or more linguistic forms have identical or largely identical meaning and it is possible to use either the one or the other variant. Variation can be free or linguistically or socially conditioned. I argue that there is an implicational relationship between complexity of the first type that is defined in terms of numbers of items and variation. Variation is a type of complexity because it implies the existence of more than one linguistic form per meaning. But not every type of complexity involves variation because complexity defined on the basis of transparency does not necessarily imply the existence of more than one form. In the following I discuss extralinguistic factors that (possibly) have an impact on socially conditioned variation and/or complexity and can lead to an increase or decrease of complexity and/or variation. I conclude with suggestions of how to further examine the relationship between complexity and variation.

Keywords: complexity, variation, transparency, quantificational approaches, extralinguistic factors

OPEN ACCESS

Edited by:

*Kilu Von Prince,
Heinrich Heine University of
Düsseldorf, Germany*

Reviewed by:

*Matthias Urban,
University of Tübingen, Germany
Sterre Leufkens,
Utrecht University, Netherlands*

***Correspondence:**

*Diana Forker
diana.forker@uni-jena.de*

Specialty section:

*This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication*

Received: 23 November 2020

Accepted: 11 March 2021

Published: 09 April 2021

Citation:

*Forker D (2021) Complexity and Its
Relation to Variation.
Front. Commun. 6:632468.
doi: 10.3389/fcomm.2021.632468*

INTRODUCTION

From time to time it is helpful to take a step back and reflect the foundations of our concepts since they represent a very important type of our tools in linguistics. COMPLEXITY and VARIATION are two such widely employed terms that at first glance do not seem to have much in common. Languages seem amazingly complex, in particular when one tries to learn foreign languages after childhood and youth. In the linguistic literature as well as in layman's understanding, complexity mostly equals with rich inflectional and derivational morphology or a large phoneme inventory. And languages seem also astonishingly varied. No one's language seems to be exactly identical to the language of other speakers, even when they are said to speak the same language. Thus, when thinking of variation within languages, dialectology or sociolinguist studies might come to one's mind.

The main goal of this paper is to point out that variation is a type of complexity and to explicate and exemplify this specific type of relation between the two concepts. I will systematically review definitions and approaches to both concepts and based on those explications show what exactly complexity and variation unites. The second aim of this paper is to review and systematize the fast-growing literature on complexity in order to show that quantificational approaches can be classified into two basic types, which are built upon two conceptually independent ways of conceptualizing complexity. In the conclusion, I will identify a number of hypotheses that can help to guide future research to deepen our understanding of correlations between extralinguistic factors and complexity or respectively variation.

DEFINING LINGUISTIC COMPLEXITY

In the literature, there are plenty of approaches and definitions of linguistic complexity. In this subsection, I will outline various approaches and definitions thereby trying to identify underlying commonalities. Subsequently I will discuss how they have been implemented in studies that examine complexity in subdomains of grammar and the lexicon.

Miestamo (2008) distinguishes between objective (or absolute) and relative (or subjective) complexity. The first type, objective complexity, is defined “in terms of the number of parts in a system.” A more complex system is constituted of more parts than a less complex system. The second type, relative complexity, can be rephrased as relative difficulty of a linguistic phenomenon for different types of language users (in particular L2 speakers and language learners, see, e.g., Kusters, 2003). Similar distinctions have been made by other linguists as well, e.g., Dahl (2004, p. 39–43) differentiates between “system complexity” (= objective complexity) and “difficulty” (= relative complexity) and Lindström (2008) between “system-based” and “user-based” complexity.

This suggests that there is a neat difference between two clearly identifiable types of complexity in language and researchers are free to decide if they want to study the one or the other. However, objective complexity defined in the way just mentioned cannot always be separated from relative complexity. If we assume the *Principle of least effort* (e.g., Zipf, 1949; Horn, 1984) and efficiency and distinctiveness pressures working in opposite directions, then objective complexity implies relative difficulty for the speaker and relative simplicity for the hearer. What is economical and efficient for speakers, namely a language as less complex as possible with ideally only one simple linguistic expression leads to infinite complexity for the hearer who has to infer all possible meanings. Vice versa, distinct expressions for every possible message means an infinite number of parts and thus a high degree of complexity for the speaker, but probably more ease for the hearer. In fact, this line of argumentation can recurrently be identified in the discussion of complexity. For example, Bisang (2009) studies under the label of “hidden complexity” analytic and isolating languages of East and mainland Southeast Asia. These languages have

comparatively little morphology such that complex expressions look formally simple. But because they express a wide range of meanings, the burden of the interpretation is carried by the hearer from whose perspective the languages can thus be categorized as complex according to Bisang (2009). This type of complexity is “hidden,” in contrast to the “overt complexity” of morphologically complex languages. Similarly, Sinnemäki (2008, 2009) bases his account of complexity in core argument marking on general principles of economy (or effectiveness) and distinctiveness, which when combined result in the principle of One–Meaning–One–Form. According to distinctiveness, one meaning is encoded by at least one form, and according to economy one meaning is encoded by no more than one form. Violations of distinctiveness and violations of economy/effectiveness can be interpreted as complexity (difficulty), respectively for the hearer and for the speaker. In other words, objective and relative complexity are in a tight relationship.

In this paper, I will follow Miestamo (2008) and use the label “objective complexity” for all conceptualizations that are based on quantification (i.e., counting items or rules or parts of items or rules), and “relative complexity” for all approaches that focus on the production, comprehension, processing and acquisition of more or less complex linguistic structures.

Karlsson et al. (2008) base their conceptualization of complexity on the classification of Rescher (1998) and reshape it for linguistics. According to their approach, linguistic complexity can be accounted for at three levels:

- the ontological level
- the epistemological or epistemic level
- the functional level

The first two levels (ontological; epistemological/epistemic) are objective in the sense of Miestamo (2008). The third functional level is processing-related and, according to the authors, refers to “cost-related differences concerning language production and comprehension” (Karlsson et al., 2008, p. ix). It is thus relative complexity in the terminology of Miestamo (2008).

The ontological level refers to which entities exist and what their relations are and thus to the language system. In other words, for an existing entity to exhibit ontological complexity means to be composed of many different interrelated components (see also Givón, 2009, p. 4). This comes close to some accounts of objective complexity that make use of the concept of “complex (adaptive) system” (e.g., Dahl, 2004; Givón, 2009; Pellegrino et al., 2009; Larsen-Freeman, 2012). The epistemological level refers to our knowledge and to be epistemologically complex means that descriptions or instructions or computations that encode our knowledge are composed of many individual steps.

Linguists have developed a number of measurements for ontological and epistemological complexity that can be classified into a few basic types (Table 1). In the following, I will first explain the main features of the approaches and then argue that all studies of objective linguistic complexity fall into one of those basic types. For that aim I will review a number of influential or exemplary works on linguistic complexity.

TABLE 1 | Approaches to linguistic complexity.

Complexity studies at the ontological level	Complexity studies at the epistemological level
<ul style="list-style-type: none"> • Counting <ul style="list-style-type: none"> ◦ Linguistic forms (McWhorter, 2007; Miestamo, 2008; Nichols, 2009; Anderson, 2015; Bentz et al., 2015) ◦ Features (Nichols, 2009; Moran and Blasi, 2014) ◦ Meanings (Gil, 2009; Matthewson, 2014) ◦ (Hierarchical) interactions and relations between forms or features (rules, regulations, constrains, etc.) (McWhorter, 2007; Nichols, 2009; Anderson, 2015; Audring, 2017) • Quantifying transparency (McWhorter, 2007; Miestamo, 2008; Sinnemäki, 2008; Nichols, 2009; Anderson, 2015; Leufkens, 2015, 2020; Audring, 2017) • Based on grammars and other published analyses and descriptions of linguistic structures or lexical items, i.e., on the language system or structure • Based on corpora, i.e., manifestations of language use (Juola, 2008; Bentz et al., 2015, 2016; Bentz, 2016; Ehret and Szmrecsanyi, 2016; Ehret, 2017) 	<ul style="list-style-type: none"> • Quantifying description length/information content by means of <ul style="list-style-type: none"> ◦ Compression algorithms (Juola, 2008; Bentz et al., 2016) ◦ Dictionary definitions (Lewis and Frank, 2016) ◦ Logical formulas (Matthewson, 2014) ◦ ...

Researchers have adopted two types of measurements for ontological complexity. The first is based on counting; the second measurement aims at quantifying the degree of transparency.

The first measurement is easier to operationalize and is therefore prevalent in typological studies that compare languages with respect to their complexity in different domains of grammar (e.g., Parkvall, 2008; Nichols, 2009; Szmrecsanyi and Kortmann, 2009). It requires to count linguistic items (phonemes, complex onsets or codas, morphemes, embedded clauses, levels of embeddings, lexemes, etc.) or, occasionally, features (phonological, features, grammatical features), or meanings (for semantic complexity) within a delimited domain. Many authors also consider interactions between linguistic items or features such as conditions, rules, and dependency relations as contributing to linguistic complexity, even though there is no unified method for the quantificational assessment of that type of complexity. The more forms, features, constructions or constraints there are (in a certain domain of grammar or overall) the more complex the language is (in that domain or in general).

The second type – quantification of the degree of transparency of linguistic forms and interactions – can be defined as any kind of violation of the principle of One-Meaning–One-Form (Dammel and Kürschner, 2008; Miestamo, 2008; Leufkens, 2015). Such violations can have various forms, e.g., syncretism and homophony, i.e., one form has more than one meaning; allomorphy and other types of variation and multiple exponence or redundancy, i.e., one meaning is expressed by more than one form; zero expression, i.e., certain meanings are not expressed at all. These violations represent lower degrees of transparency or regularity and thus higher complexity than simple one-to-one

form-meaning relationships. The more the formal coding (in a certain domain of grammar or overall) adheres to transparency, the less complex it is. The quantification of the degree of transparency is more difficult than just counting items because it requires the objective rating of the different types of violations (if one does not simply want to count every irregular item).

Complexity at the epistemological level assumes that linguistic forms and structures can be adequately articulated in descriptions, instructions or computations that represent our knowledge of them. It is mainly quantified and measured in terms of description length and/or (un)predictability by means of two types of measurement that originate from information theory: *Shannon entropy* and *Kolmogorov complexity* (Juola, 1998, 2008; Dahl, 2004, p. 9–10, 21, 2009, p. 51; Fenk-Oczlon and Fenk, 2008; Miestamo, 2008; Ehret and Szmrecsanyi, 2016). In information theory, maximal randomness and unpredictability means maximal information content. A maximal unpredictable message requires the longest possible description (which is basically the length of the message itself). An alternative but related approach measures only the length of descriptions of structured patterns (e.g., grammatical rules) and therefore quantifies the degree of regularity. Other more informal ways of resorting to description length are, e.g., counting the length of definitions of lexical items in dictionaries (Lewis, 2016; Lewis and Frank, 2016) or of logical formulas used in formal semantics (Matthewson, 2014) as a proxy of the semantic complexity of linguistic expressions. The latter two measurements involve counting, but in contrast to directly counting parts of the language system they count the length of representations of specific parts of the system.

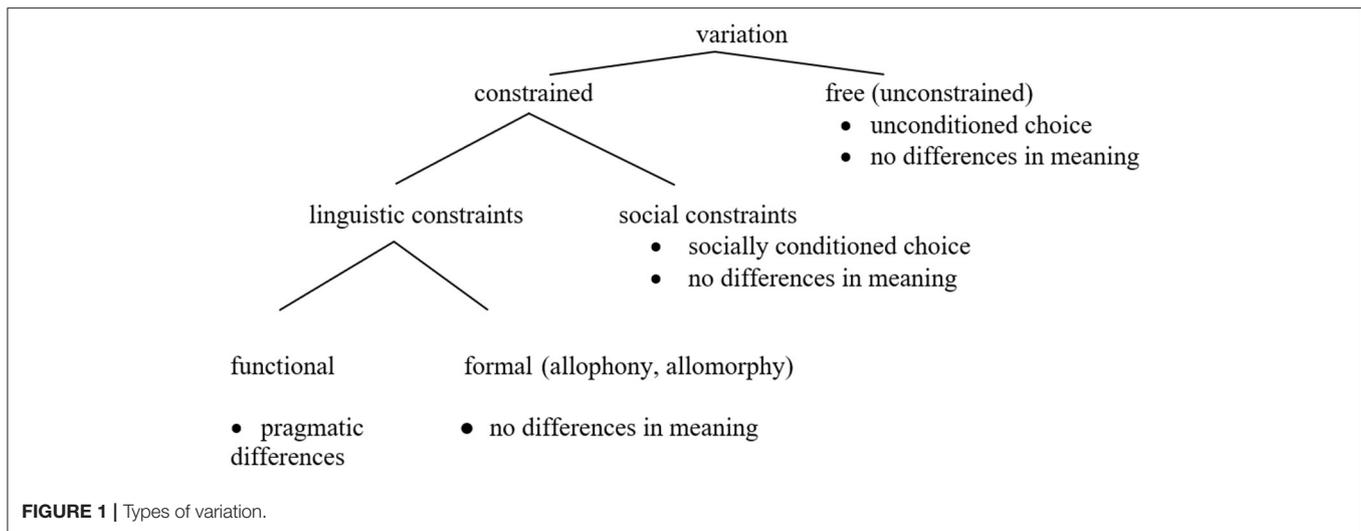
In general, there are comparably few studies that take the epistemological level seriously and apply it, in particular to instantiations of language use in the form of natural texts. Exemplary studies include Juola (1998, 2008), Bentz et al. (2016), Ehret and Szmrecsanyi (2016) and Ehret (2017), which employ text corpora as data basis and study morphological and morphosyntactic complexity. By contrast, the larger part of the studies, in particular with respect to phonological complexity, is based on pre-fabricated linguistic analyses in the form of grammatical descriptions and to some extent also dictionaries (for semantic complexity).

Linguistic complexity and complexity of individual languages or groups of languages has been investigated with respect to all grammatical domains, namely phonology, morphology, syntax and semantics, but to different extents and partially within very heterogenous approaches. Some researchers have focused on one domain only. Others have attempted to compare languages based on more than one domain (usually phonology and morphosyntax).

LINGUISTIC VARIATION

In this section, I will define the concept of variation, discuss different types of variation and methods how to study them.

In a very general sense, the terms “variation” and “variants” can be defined as referring to a situation when two (or more)



linguistic items (i.e., forms) have identical or largely identical meaning and it is possible to use either the one or the other variant to express the same semantics, but possibly with different pragmatic functions. Another type of variation concerns frequency: one and the same linguistic item can be used more or less frequently. **Figure 1** displays the different types of linguistic variation. The term “meaning” in this schema refers to the linguistic meaning in the sense of semantics, not to social or otherwise non-linguistic forms of meaning.

Variation can be free, which means that speakers always have the choice between one or the other variant with no difference in linguistic meaning or social meaning between the two variants. A very simple example are the German words *Sofa* and *Couch* that have the same meaning and whose use is not constrained by regional or social provenance of the speaker.

When variation is constrained, the constraints are either inherent to the language and thus linguistic or they are extralinguistic. In the case of linguistic constraints, the choice of the speaker is conditioned by, e.g., subtle differences in pragmatics as it is possible for alternative constituent orders in German, e.g., *Ich geb dir das Buch.* vs. *Dir geb ich das Buch* vs. *Das Buch geb ich dir* (“I give you the book.”). Or the constraints can be formal, e.g., regulated by phonological/phonetic properties or be lexical idiosyncrasies. If formal constraints exclude each other (complementary distribution), we speak of allophony or allomorphy. Speakers have no choice and the use of the variants is predictable. If the constraints that regulate the use of the variants are social (i.e., extralinguistic), then speakers have, in principle, a choice. This type of variation is at the heart of variationist sociolinguistic studies. Simply speaking, sociolinguistic variation refers to “alternative ways of ‘saying the same thing.’” (Labov, 1969, p. 738). According to Nagy and Meyerhoff (2008, p. 5) “the quantitative analysis of variation requires the researcher to first identify variants that are semantically (or, some would argue, functionally) equivalent, and then explore the (linguistic or social) constraints on the distribution of those variants.” Examples of variation within various subdomains of grammar are, e.g., alternations in the pronunciation of the phoneme /ç/ as

[ʃ] or [ç] in certain varieties of German (Jannedy and Weirich, 2014), variants of phonemes such as aspirated vs. unaspirated voiceless stops in English, variation in the form of the English gerund *read-in’* vs. *read-ing*, variation in the use of definite articles vs. possessive pronouns (e.g., *the hand* vs. *my hand*) in doctor-patient interactions in English (Fasold and Preston, 2007), or the English dative alternation.

It is important to keep in mind that variation is usually conditioned not just by one type of constraint, but by several constraints, and that the conditions can change. Thus, the pronunciation variants [ʃ] and [ç] are allophones in some dialects of central Germany, and the allophony at least partly results from a merger of the phonemes /ç/ and /ʃ/ to /ç/ (Jannedy and Weirich, 2014). The use of the variants, in particular [ʃ], has become a salient phonetic feature of Hood German - a variety spoken and associated with young people belonging to urban multiethnic networks - and thus socially conditioned. In the last years, researchers have observed that the variants are becoming less and less associated with a particular social group and instead variability becomes the norm for all speakers.

What counts as variants of one and the same linguistic variable can be problematic due to the theoretical background of the linguists and the concomitant linguistic analysis of the variable-variants-complex, but also because of the alleged functional equivalence or origin of the variants.

For instance, Cornips and Corrigan (2005, p. 9) notice that mismatches in number agreement of preverbal subjects (*When the grapes was/were in season*) are treated on a par with mismatches in expletive *there*-construction with post-verbal subjects (*There was/were two priests [who] lived there*) as variants of one and the same variable by variationist linguists. By contrast, for generativists the two constructions are not only different, but remote because of their diverging syntactic behavior.

Buchstaller (2009) points out that beyond the level of phonetics and phonology, the question of semantic or functional equivalence is far from being trivial. If we adopt the definition of morphemes as smallest meaningful elements, an alternation between two morphemes such as the definite article and a

possessive pronoun in a noun phrase necessarily correlates with a semantic alternation. The same reasoning applies to syntactic variation. Cheshire (2005, p. 85) states that “A tacit consensus seems to be that the condition of strict semantic equivalence can be relaxed for syntactic variables, so that a variable can be set up on the basis of an equivalence in discourse function.”

Similarly, Nagy and Meyerhoff (2008, p. 5) note that linguistic variants can come from more than one language. In such cases, functional or semantic equivalence of the variants is also problematic. Therefore, multilingual communities represent special challenges to variationist approaches.

We can distinguish between internal and external sources or causes (and thus explanations) for variation in language (Nagy and Meyerhoff, 2008). External sources for variation are language contact, i.e., the impact of one variety upon another, spatial, sociocultural, and biological factors. The latter are general biological characteristics and/or cognitive capacities of human beings that result in constraints on language/speech production, perception and processing. Internal factors are often called “linguistic” because they are assumed to pertain to the language or linguistic system. Allophones or allomorphy are examples of linguistic or internal variation. Variation that has been explained by resorting to concepts such as animacy, definiteness, specificity, information structure and the like is also classified as “internal” or “linguistic” (e.g., Fasold and Preston, 2007).

There is a principled distinction between intra-speaker vs. inter-speaker variation, i.e., variation at the level of the individual language user vs. variation at the level of a group of speakers. Intra-speaker variation is partly a matter of sociocultural circumstances and partly of individual biological (i.e., cognitive and other) properties (Dabrowska, 2015a) and because of the latter can be related to relative complexity.

Variation can be studied at the synchronic as well as at the diachronic level. Synchronic variation can be an indicator of an ongoing change and thus of diachronic variation, but it can also be (relatively) stable over longer periods of time. Variation can be quantified and measured in a way comparable to quantificational complexity measures (Table 1). This point will be further elaborated in Section Studying Variation vs. Studying Complexity. Quantificational approaches to variation are largely focused on socially conditioned variation, for which there are standard methodological tools that basically consist in counting items (distinct variants of one and the same variable) and their frequency of usage patterns. One also finds quantificational studies of semantically conditioned variation, e.g., Bresnan and Ford (2010) on the dative alternation. To my knowledge, there are no approaches to variation at the epistemological level making use of information-theoretic methods as they are used in complexity studies.

THE RELATIONSHIP BETWEEN COMPLEXITY AND VARIATION

If we have another look at the classifications in Table 1 and in Figure 1, it becomes clear that we can draw connections between variation and the ontological level of complexity, in

particular with respect to transparency and the principle of One-Meaning–One-Form. Variation – at least in the most common understanding – refers to the formal aspect of language because it rests on the availability of two or more different forms with normally roughly identical linguistic meaning. Therefore, variation represents a violation of the One-Meaning–One-Form principle because one meaning is expressed by more than one form. And in this sense variation can also be related to complexity understood in terms of numbers of items (“counting” in Table 1): the more forms there are the more variation and complexity there is. In other words, variation presupposes a certain type of objective complexity in the ontological sense as a property of a language (measured at the ontological or epistemological level). Or, to put it the other way around, only if at least two forms that express the same meaning are available and thus we deal with a more complex situation than in the simple One-Meaning–One-Form case, speakers have a choice between two variants. This means that there is an implicational relationship between complexity and variation: variation is a type of complexity, but not every type of complexity involves variation. Variation is a hyponym and a subordinate concept to complexity. The relation does not work the other way around, i.e., complexity does not presuppose variation because not every form of complexity consists in violations of the One-Meaning–One-Form principle. A grammatical rule whose application is restricted by many conditions is more complex than a rule that can be applied without exceptions. Any types of irregularities contribute to complexity, but not (necessarily) to variation.

In the literature on complexity and variation one can find statements that point out a relation between the two concepts, but they do not claim that it is a type-of relationship. Variation in the form of allophony or allomorphy has been claimed to contribute to linguistic complexity (e.g., McWhorter, 2007; Nichols, 2009; Szmrecsanyi and Kortmann, 2009; Anderson, 2015). Ohala (2009, p. 54) argues that phonetic variation must be included when measuring phonological complexity, because phonetic variants of segments are part of speakers’ and hearers’ knowledge of the language. In a similar vein Maddieson (2009, p. 100) maintains that free variation implies complexity: “languages for which the patterns of variation in the phonology are more “transparent” are simpler than those for which the variations are more arbitrary.”

For Braunmüller (2016) complexity naturally and logically results from spatially and socially conditioned variation. His definition of complexity differs from the one presented in Section Defining Linguistic Complexity and is rather reminiscent of variation: “Complexity emerges whenever a grammatical category or structure is represented by more than one category, form, or construction with approximately the same meaning” (Braunmüller, 2016, p. 51).

Szmrecsanyi (2015) discusses what he calls “variational complexity,” which he defines as “the extent to which choosing between linguistic variants is subject to restrictions.” The more constraints there are on variation and the more interaction between the constraints, the larger is the ontological complexity. At the same time the degree of epistemological complexity is also higher because more description is required, and he suggests

that the degree of relative complexity in terms of difficulty for language acquisition is larger as well. Furthermore, since variation is not just about the language system and its parts, but also about speaker-made language choices and frequency patterns, the concept of variational complexity also extends to usage (“procedural complexity” in his terms). A comparable approach to complexity that focuses on the use of linguistic items instead of their simple existence can be found in a paper by Van den Broeck (1977). Instead of analyzing why a construction or language or another linguistic item IS more complex than another he points out that linguists should ask why certain speakers use more complex forms than others or why one and the same speaker uses more complex forms in situation X than s/he uses in situation Y. Van den Broeck (1977, p. 164–165) suggests a number of possible answers regarding the functional value of more complex syntactic constructions. Because of iconicity, it could be the case that more complex topics are expressed by more complex syntactic constructions. From the perspective of shared knowledge and experience it would be conceivable that interlocutors who know each other less well tend to be more explicit and use more complex syntactic constructions. From the perspective of style, van den Broeck hypothesizes that certain complex constructions could be *en vogue* similar to lexical items. After arguing against the three possible explanations he states that “the use of more complicated forms is an act of ‘conspicuous ostentation’, a means of display, a marker of social distance.” He further proposes a relationship between variation in phonology and syntax and the formality of situations: in formal situations speakers use a larger variety of syntactic constructions but a smaller variety of phonological variants than in informal situations where the relation is the opposite. If we replace variation with complexity the hypothesis can be rephrased: we expect more syntactic complexity and less phonological complexity in formal situations in which speakers carefully monitor their speech than in informal situations.

In the following section, I will point out parallels and differences in the study of socially conditioned variation, in particular with respect to extralinguistic constraints and diachrony. I will use the term “variation” instead of “socially conditioned variation,” but concentrate only on this type and neglect the other types given in **Figure 1**.

STUDYING VARIATION VS. STUDYING COMPLEXITY

In theory, we can study variation and complexity at the level of the individual speaker (intra-speaker variation), in a speech community of whatever size, in other words within a language (inter-speaker variation), and also across different languages (cross-linguistically). With respect to variation, the group or community level is prevalent, but variation in the speech of individual speakers may also constitute the object of inquiry. At both levels, quantificational methods play a major role for determining the extent of variation and identifying correlations with linguistic and extralinguistic factors. Cross-linguistic studies of variation are absent or rare and sociolinguistic typology is a

relatively new field. By contrast, objective complexity is usually studied at the level of individual languages (or grammatical domains of individual languages) and regularly compared across languages (i.e., across speech communities), but not examined at the level of the individual speaker. We know from a few studies that there are individual differences in our linguistic abilities (e.g., Chipere, 2009; Dabrowska, 2015a; Petré and Anthonissen, 2020). These differences between specific speakers and their grammars could, in principle, be examined at the ontological level by counting parts of their language systems or by quantifying the transparency of the constructions that they use. Both complexity studies and variationist studies make use of quantificational methods and search for correlations with linguistic and extralinguistic factors. Variationist studies basically count items, whereas complexity studies employ a larger range of tools (counting items, feature, and interactions, determining transparency and approaches from information theory based on description length, entropy, etc., **Table 1**).

The question of how or where variation should be explained has repeatedly been debated, and there are basically two opposing answers: within the linguistic system by means of optional rules, different rule orders or the like or outside of the linguistic system by means of social factors. In the first case, variation is assumed to be an inherent property of grammars. In the second case variation can, for instance, be explained by recourse to separate grammars between which speakers can choose analogously to bilingual speakers who might switch between two different languages. In contrast, complexity as a property of certain grammatical domains does not imply choices because grammaticalized meaning distinctions such as gender, which adds complexity to the languages that have it, are obligatory (Nichols, 2019).¹ In languages with gender systems speaker normally do not have the choice to express or not express the gender of referents.

Variation and complexity also differ with respect to their functions. Variation has repercussions at the level of language use because speakers have a choice. As variationist sociolinguists have shown over and over again, socially conditioned variants are loaded with extralinguistic meaning and thus serve social functions for speakers and hearers. By contrast, the function of complexity, if there is any, can be viewed as enhancing distinctiveness, which is supposed to help the hearer (section Defining Linguistic Complexity).

Next, I will discuss extralinguistic constraints on variation and complexity and in particular the question whether particular findings concerning complexity can be replicated for variation or vice versa. The factors are interrelated, which should be kept in mind even though I provide them here in the form of a table (**Table 2**). They can be divided into factors that depend on the individual speaker and factors that operate at the level of groups of various kinds (clans, networks, speech communities, states, etc.). Some factors operate at both levels

¹ However, as one reviewer pointed out, the development of obligatory grammatical markers and thus grammaticalization reflects linguistic behavior and thus linguistic choices of past speakers. In other words, today’s grammatical markers are obligatory but they go back to certain selections of earlier generations of speakers.

TABLE 2 | Extralinguistic factors possibly impacting on variation and/or complexity.

Individual level	Group/society/community/state level
Age	Community size
Gender	Social organization (individualistic vs. collectivistic)
Cognitive abilities, working memory, etc.	Network density (dense vs. loose)
Education and profession	Standardization and literacy development
Attitude, ...	Proportion of L2 learners
Geographical location	
Functional domains of language use	
Language contact and bilingualism/multilingualism	

(e.g., bilingualism is an individual property but can also be a feature of an entire community).

Starting with the impact of individual factors on complexity we can say that these studies fall into the scope of relative complexity. They are examined in psycholinguistics and in applied linguistics and encompass production and comprehension studies with a focus on syntax (see Friedrich, 2019, p. 68–123 for a summary of recent studies; Jin et al., 2020). Complexity measures most frequently used are sentence length, mean length of utterance in morphemes, structure in terms of types and number of embedded clauses and level of embedding (Cheung and Kemper, 1992; Kyle and Crossley, 2018), but also semantic content defined as idea density or propositional density.

The age factor has been investigated in many studies with unclear results. However, it seems that elderly speakers lose some linguistic capacities but because they gain others there can be compensatory effects (Friedrich, 2019, p. 125–132). With respect to vocabulary there is obviously an increase with age and according to a recent study the peak performance seems to occur as late as late as in the 60's (Hartshorne and Germine, 2015). Gender does not seem to have an effect and education shows perhaps a small positive correlation with an increase in linguistic complexity (except, of course, for vocabulary size that correlates with education, social class, and ethnic background; Farkas and Beron, 2004; Friedrich, 2019, p. 120). Furthermore, working memory has an important impact on language production and comprehension and cannot be easily separated from linguistic abilities (Chipere, 2009; Dabrowska, 2015a).

By contrast, sociolinguists have repeatedly found correlations between particular variants and individual factors such as age, gender, education, profession, etc. For instance, many studies have shown that teenagers are more innovative than other age groups (e.g., Tagliamonte and D'Arcy, 2009) and that at least in western societies females adhere more to the standard than males for certain linguistic variables while for other variables they are more innovative than men (e.g., Meyerhoff, 2006, p. 207–222). Speakers with higher education show less variation than speakers with lower levels of education because their linguistic skills have been shaped by many years of formal instruction in one particular language variety – the standard language (Dabrowska, 2015a).

Continuing with the impact of community-level factors on complexity, community size in combination with network density has especially been in the focus of research. It has been reported that complex morphology is predominantly found in small languages with dense networks because in intergenerational language transmission it is easier to ensure the preservation of complexity within smaller groups than within larger communities with loose networks (Trudgill, 2009, 2011; Lupyán and Dale, 2010). “[S]mall, tightly-knit communities are more able to encourage the preservation of norms, and the continued adherence to norms, from one generation to another, however complex they may be” (Trudgill, 2009, p. 102). Another claim by the same author that rather goes in the opposite direction is that small communities “will have large amounts of shared information in common and will therefore be able to tolerate lower degrees of linguistic redundancy of certain types” (Trudgill, 2004, p. 306), which can be exemplified by small languages with small phoneme inventories such as the Polynesian languages.² This statement plainly contradicts the previous one on morphology because it declares that small communities tend to have less complex languages. In fact, research concerning phonological complexity has not produced clear and consistent results regarding the role of community size. A number of studies have found the opposite of Trudgill's claim, namely a positive correlation between community size and size of the phoneme inventory (see Nettle, 2012 and references therein), but Moran et al. (2012) argue against those findings. With respect to the lexicon it seems that the picture is rather clear: bigger languages with standardized forms, developed literacy and covering all functional domains have a larger lexicon (Reali et al., 2018), but there are no studies that consider other types of semantic complexity. Furthermore, it has repeatedly been stated that standard, mostly written varieties are more complex than spoken, vernacular varieties with respect to morphosyntactic properties such as complex and subordinate clause formation exactly because of the written mode (e.g., Dahl, 2009; Szmrecsanyi and Kortmann, 2009; Dabrowska, 2015b; Baechler, 2016, p. 17; Braunnüller, 2016).

An example of a study that examines the impact of geographical location in terms of latitude on complexity is Moran and Blasi (2014). They find a positive correlation between latitude and the number of obstruents and latitude and syllable structure, which means that the further to the north a language is spoken, the more complex is its obstruent system and syllable structure.

²Leufkens (2020) specifically focuses on four types of syntagmatic morphosyntactic redundancy and concludes that we have to distinguish between two kinds of redundancy, which differ in function and diachronic origin. The first type is called accidental redundancy and it arises in the case of obligatory morphosyntactic markers. Broadly speaking, this type enhances the successful transmission of messages because it repeats information and thus improves saliency and preciseness. The second type is called purposeful redundancy and is found with optional markers that are used for pragmatic effects such as emphasis. It would be worth to check if Leufkens' 50 language sample shows correlation with community size. We could hypothesize that the larger languages show higher levels of accidental redundancy and smaller languages show higher levels of purposeful redundancy. In the first case speakers need to be more precise because they share less knowledge. In the second case extravagant pragmatic effects may get more attention in smaller communities.

Other exemplary studies that claim environmental effects on complexity measures are Everett (2013) on ejectives, Everett et al. (2016) on tone, Everett (2017) on vowel richness, and Bentz (2016) on languages just above the equator that exhibit lower complexity than languages further away from the equator when compared by means of information-theoretic complexity measured in corpora.

As regards language contact and bilingualism, researchers have found that a high rate of child bilingualism is often a driving force for complexification (e.g., Nichols, 1992, p. 192–195), whereas high numbers of second language learners rather lead to simplification (e.g., Szmrecsanyi and Kortmann, 2009; Trudgill, 2009).

Among the community-level factors that might impact variation, geographical location and network density are the most frequently researched aspects. The impact of language contact has also been considered, in particular in situations of language shift under attrition, which have been shown to lead to a “larger than usual” extent of variation (Cook, 1989; Dorian, 1989; Babel, 2009). Although there are, to my knowledge, no studies that have compared the amount of variation between languages and tested correlations with community size, it is reasonable to hypothesize that in larger speech communities there is more variability and thus more variation simply due to the larger number of individual speakers. A study by Atkinson et al. (2015) tested whether languages spoken by a bigger community which therefore display a larger amount of variability undergo simplification processes in their morphology because faithful cross-generational transfer is more difficult, but they did not find evidence for this hypothesis. The hypothesis reminds of well-known processes of dialect leveling by which variation within dialects and between dialects, in particular in relationship to the standard variety, is reduced through convergence, assimilation and mixture (e.g., Hinskens, 1998; Meyerhoff, 2006, p. 239–240; Noglo, 2009). In other words, a large amount of variation can, in fact, lead to simplification.

Dialect leveling is a type of diachronic change, and thus leads us to the discussion of the diachronic dimension of studying variation and/or complexity. Givón (2009, p. 8) notes that (syntactic) complexity plausibly arises by means of a process of synthesis or combination (as opposed to a theoretically possible opposite process of decomposition and reanalysis): simple linguistic items are combined into complex items. Dahl (2004, p. 293) concludes that under “normal ecolinguistic conditions” up to a certain point, languages tend to become more complex rather than less complex. Dialect leveling does not represent “normal ecolinguistic conditions” but rather involves high-contact situations of adult speakers which, as was said above, typically leads to simplification.

SUGGESTIONS FOR FUTURE DIRECTIONS OF RESEARCH AND CONCLUDING REMARKS

In this paper I have discussed linguistic complexity, in particular the concepts of OBJECTIVE COMPLEXITY and VARIATION and links between them. I have shown that studies of objective

complexity can be classified into two basic types, namely counting various kinds of items and determining transparency. I have argued that there is an implicational relationship between complexity and variation: variation is a type of complexity, but not every type of complexity involves variation. I have then sketched the main directions of research and methodological approaches when studying complexity vs. variation and pointed out similarities and differences.

Future explorations of complexity could profit from a collaboration with variationist sociolinguists. Vice versa, researchers within the variationist paradigm can open up their perspective and extend their methodological tools and research questions by taking into consideration complexity studies. Research on English can serve as an example to illustrate overlapping points between complexity and variationist endeavors. Judging from the countless sociolinguistic studies on varieties of English it seems that English exhibits a rather large degree of variation. At the same time, English is normally classified as being not very complex (e.g., Juola, 2008; Parkvall, 2008), which has repeatedly been explained by large numbers of L2 speakers. Thus, a high degree of variation goes hand in hand with a low degree of complexity. But this does not necessarily have to be the case for other languages because the various extralinguistic influencing factors can work in different directions.

In particular, investigations of extralinguistic factors that play a role in explaining causes of complexity and variation can be a fruitful area of overlap for future research. Community size in combination with network density is probably the most commonly explored extralinguistic impact factor in complexity studies. The latter factor is well-known in variationist approaches, but mere community size is normally not considered. There are a number of dependent factors that might increase or reduce variation and/or complexity that are not found to the same extent in large vs. small speech communities and I will propose possible correlations that could be examined in the future.

A bigger speech community consists of more speakers and thus of a potentially bigger pool for linguistic innovators that introduce and propagate new variants and therefore more variation than in a small community. At the same time, small communities with dense social networks might exercise more control over their members and thus suppress variation [see the quote by Trudgill (2009) above about the adherence and preservation of norms]. However, in a larger community, innovations are probably less visible and, at least theoretically, might have a bigger pay-off in smaller communities.

Standardization aims at imposing a homogenous variety. It is normally planned and enforced by an official language policy, which, in turn, is usually restricted to larger national languages. Therefore, we can hypothesize that standardization leads to less variation in big communities.

The impact of standardization on complexity could be contradictory and go in opposite directions. Comparable to variation, standardization might reduce complexity in those cases in which standard varieties have been created by processes of dialect leveling or language planning (e.g., Byron,

1976 on the standardization of Albanian; Trudgill, 2009 on English; Dabrowska, 2015a). On the other side, standardization predominantly effects formal styles and the written mode, for which an increase in syntactic complexity as opposed to oral, vernacular varieties is normally attested (Dabrowska, 2015b).

Different types of language contact phenomena and bilingualism/multilingualism can be conjectured to lead to drifts in opposite directions. Speakers of smaller languages often have a greater need to know other languages and this knowledge might influence their own language use and thus be a source for variation. However, the use of a large language as lingua franca may also lead to diglossic situations and relatively clear functional separation such that the minority language largely remains untouched by the lingua franca and serves as a clear identity marker of the minority speech community (Braunmüller, 2016).

In larger speech communities the proportion of L2 speakers is often higher, which might itself be a source for variation. We also know from the work of Trudgill (2011) and others that a high proportion of L2 speakers can lead to considerable simplification of big languages. Braunmüller (2016, p. 49) convincingly maintains that such a situation also leads to more variation because L2 speakers introduce linguistic innovations based on transfer and imperfect learning.

REFERENCES

- Anderson, S. R. (2015). "Dimensions of morphological complexity," in *Understanding and Measuring Morphological Complexity*, eds M. Baerman, D. Brown and G. G. Corbett (Oxford: Oxford University Press), 11–26. doi: 10.1093/acprof:oso/9780198723769.003.0002
- Atkinson, M., Kirby, S., and Smith, K. (2015). Speaker input variability does not explain why larger populations have simpler languages. *PLoS ONE* 10:e0129463. doi: 10.1371/journal.pone.0129463
- Audring, J. (2017). Calibrating complexity: how complex is a gender system? *Lang. Sci.* 60, 53–68. doi: 10.1016/j.langsci.2016.09.003
- Babel, M. (2009). "The phonetic and phonological effects of obsolescence in Northern Paiute," in *Variation in Indigenous Minority Languages*, eds J. Stanford, and D. Preston (Amsterdam: Benjamins), 23–45. doi: 10.1075/impact.25.03bab
- Baechler, R. (2016). "Inflectional complexity of nouns, adjectives and articles in closely related (non-)isolated varieties," in *Complexity, Isolation, and Variation*, eds R. Baechler, and G. Seiler (Berlin: de Gruyter), 15–46. doi: 10.1515/9783110348965-002
- Bentz, C. (2016). "The low-complexity-belt: evidence for large-scale language contact in human prehistory?" in *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*, eds S.G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, and T. Verhoef. Available online at: <http://evolang.org/neworleans/papers/93.html> (accessed March 23, 2021).
- Bentz, C., Soldatova, T., Kopenig, A., and Samardžić, T. (2016). "A comparison between morphological complexity measures: typological data vs. language corpora," in *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity (CLALC)* (Osaka), 142–153.
- Bentz, C., Verkerk, A., Kiela, D., Hill, F., and Buttery, P. (2015). Adaptive communication: languages with more non-native speakers tend to have fewer word forms. *PLoS ONE* 10:e0128254. doi: 10.1371/journal.pone.0128254
- Bisang, W. (2009). "On the evolution of complexity – Sometimes less is more in East and Mainland Southeast Asia," in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 34–49.
- In addition to community size (which is comparatively easy to estimate) and network density (which is more difficult to define and establish), it could be worth to examine the impact of social organization in the sense of a broad classification of societies into individualistic vs. collectivistic. Individualistic societies could be expected to exhibit a greater degree of variation, but complexity is perhaps better preserved in collectivist societies.
- There are also open questions regarding the diachronic dimension. If variation is a type of complexity, then an increase in variation immediately means an increase in complexity. But can we also find a correlation in the other direction, i.e., is a growth in complexity always accompanied by a growth in variation?

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I am grateful to Stefan Schnell for helpful suggestions and discussions. I also want to thank two reviewers for their comments and recommendations.

- Braunmüller, K. (2016). "On the origins of complexity: evidence from Germanic," in *Complexity, Isolation, and Variation*, eds R. Baechler and G. Seiler (Berlin: de Gruyter), 47–69. doi: 10.1515/9783110348965-003
- Bresnan, J., and Ford, M. (2010). Predicting syntax: processing dative constructions in American and Australian varieties of English. *Language* 86, 168–213. doi: 10.1353/lan.0.0189
- Buchstaller, I. (2009). The quantitative analysis of morphosyntactic variation: constructing and quantifying the denominator. *Lang. Linguist. Compass* 3.4, 1010–1033. doi: 10.1111/j.1749-818X.2009.00142.x
- Byron, J. L. (1976). *Selection Among Alternates in Language Standardization: The Case of Albanian*. The Hague: Mouton. doi: 10.1515/9783110815931
- Cheshire, J. (2005). "Syntactic variation and spoken language," in *Syntax and Variation: Reconciling the Biological and the Social*, eds L. Cornips and K. P. Corrigan (Amsterdam: Benjamins), 81–106. doi: 10.1075/cilt.265.05che
- Cheung, H., and Kemper, S. (1992). Competing complexity metrics and adults' production of complex sentences. *Appl. Psycholinguist.* 13, 53–76. doi: 10.1017/S0142716400005427
- Chipere, N. (2009). "Individual differences in processing complex grammatical structures," in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil and P. Trudgill (Oxford: Oxford University Press), 178–191.
- Cook, E. (1989). Is phonology going haywire in dying languages? Phonological variations in Chipewyan and Sarcee. *Lang. Soc.* 18, 235–255. doi: 10.1017/S0047404500013488
- Cornips, L., and Corrigan, K. (2005). "Toward an integrated approach to syntactic variation: a retrospective and prospective synopsis," in *Syntax and Variation: Reconciling the Biological and the Social*, eds L. Cornips, and K. P. Corrigan (Amsterdam: Benjamins), 1–27. doi: 10.1075/cilt.265.01cor
- Dabrowska, E. (2015a). "Individual differences in grammatical knowledge," in *Handbook of Cognitive Linguistics*, eds E. Dabrowska, and D. Divjak (Berlin: De Gruyter), 649–667. doi: 10.1515/9783110292022
- Dabrowska, E. (2015b). "Language in the mind and in the community," in *Change of Paradigms—New Paradoxes: Recontextualizing Language and Linguistics*, eds J. Daems, E. Zenner, K. Heylen, D. Speelman, and H. Cuyckens (Berlin: De Gruyter), 221–235.
- Dahl, Ö. (2004). *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: Benjamins. doi: 10.1075/slcs.71

- Dahl, Ö. (2009). "Testing the assumption of complexity invariance: the case of Elfdalian and Swedish," in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 50–63.
- Dammel, A., and Kürschner, S. (2008). "Complexity in nominal plural allomorphy: A contrastive survey of ten Germanic languages," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki and F. Karlsson (Amsterdam: Benjamins), 243–262. doi: 10.1075/slcs.94.15dam
- Dorian, N. (Ed.). (1989). *Investigating Obsolescence*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511620997
- Ehret, K. (2017). *An information-theoretic approach to language complexity: Variation in naturalistic corpora* (Doctoral dissertation). Breisgau: University of Freiburg. doi: 10.1515/cltt-2018-0033
- Ehret, K., and Szmrecsanyi, B. (2016). "An information-theoretic approach to assess linguistic complexity," in *Complexity, Isolation, and Variation*, eds R. Baechler and G. Seiler (Berlin: de Gruyter), 71–94. doi: 10.1515/9783110348965-004
- Everett, C. (2013). Evidence for direct geographic influences on linguistic sounds: the case of ejectives. *PLoS ONE* 8:e65275. doi: 10.1371/journal.pone.0065275
- Everett, C. (2017). Languages in drier climates use fewer vowels. *Front. Psychol.* 8:1285. doi: 10.3389/fpsyg.2017.01285
- Everett, C., Blasi, D. E., and Roberts, S. G. (2016). Language evolution and climate: the case of desiccation and tone. *J. Lang. Evolut.* 1, 33–46. doi: 10.1093/jole/lzv004
- Farkas, G., and Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: differences by class and race. *Soc. Sci. Res.* 33, 464–497. doi: 10.1016/j.ssresearch.2003.08.001
- Fasold, R., and Preston, D. (2007). "The psycholinguistic unity of inherent variability: old occam whips out his razor," in *Sociolinguistic Variation*, eds R. Bayley and C. Lucas (New York, NY: Cambridge University Press), 45–69. doi: 10.1017/CBO9780511619496.004
- Fenk-Oczlon, G., and Fenk, A. (2008). "Complexity trade-offs between the subsystems of language," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki, and F. Karlsson (Amsterdam: Benjamins), 43–65. doi: 10.1075/slcs.94.05fen
- Friedrich, L. (2019). *Sprachliche Komplexität zwischen kognitiven Veränderungen, Individualität und Prädiktion* (Doctoral dissertation). Mainz: University of Mainz.
- Gil, D. (2009). "How much grammar does it take to sail a boat?," in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 19–33.
- Givón, T. (2009). *The Genesis of Syntactic Complexity*. Amsterdam: Benjamins. doi: 10.1075/z.146
- Hartshorne, J. K., and Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychol. Sci.* 26.4, 433–443. doi: 10.1177/0956797614567339
- Hinskens, F. (1998). Dialect levelling: a two-dimensional process. *Folia Linguist.* 32, 35–52. doi: 10.1515/flin.1998.32.1-2.35
- Horn, L. (1984). "Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature," in *Meaning, Form, and Use in Context*, ed D. Schiffrin (Washington: Georgetown University Press), 11–42.
- Jannedy, S., and Weirich, M. (2014). Perceptual divergence in an urban setting: category in-stability of the palatal fricative in Berlin. *Lab. Phonol.* 5, 91–122. doi: 10.1515/lp-2014-0005
- Jin, T., Lu, X., and Ni, J. (2020). Syntactic complexity in adapted teaching materials: differences among grade levels and implications for benchmarking. *Modern Lang. J.* 104, 192–208. doi: 10.1111/modl.12622
- Juola, P. (1998). Measuring linguistic complexity: the morphological tier. *J. Quant. Linguist.* 5, 206–213. doi: 10.1080/09296179808590128
- Juola, P. (2008). "Assessing linguistic complexity," in *Language Complexity: Typology, contact, change*, eds M. Miestamo, K. Sinnemäki and F. Karlsson (Amsterdam: Benjamins), 89–108. doi: 10.1075/slcs.94.07juo
- Karlsson, F., Miestamo, M., and Sinnemäki, K. (2008). "Introduction: the problem of language complexity," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki and F. Karlsson (Amsterdam: Benjamins), i–xiv. doi: 10.1075/slcs.94.01kar
- Kusters, W. (2003). *Linguistic complexity: the influence of social change on verbal inflection* (Ph.D. Dissertation), University of Leiden; LOT, Utrecht.
- Kyle, K., and Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *Modern Lang. J.* 102, 333–349. doi: 10.1111/modl.12468
- Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45, 715–762. doi: 10.2307/412333
- Larsen-Freeman, D. (2012). "Preface: a closer look," in *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact*, eds B. Kortmann, and B. Szmrecsanyi (Berlin: de Gruyter), 1–5. doi: 10.1515/9783110229226.1
- Leufkens, S. (2015). *Transparency in language: A typological study* (University of Amsterdam PhD dissertation). Utrecht: LOT.
- Leufkens, S. (2020). A functionalist typology of redundancy. *Revista Da Abralin* 19, 79–103. doi: 10.25189/rabralin.v19i3.1722
- Lewis, M. L. (2016). *Conceptual complexity and the evolution of the lexicon* (Doctoral dissertation). Stanford, CA: Stanford University. doi: 10.31237/osf.io/6c9n2
- Lewis, M. L., and Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition* 153, 182–195. doi: 10.1016/j.cognition.2016.04.003
- Lindström, E. (2008). "Language complexity and interlinguistic difficulty," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki and F. Karlsson (Amsterdam: Benjamins), 217–242. doi: 10.1075/slcs.94.14lin
- Lupyan, G., and Dale, R. (2010). Language structure is partly determined by social structure. *PLoS ONE* 5:e8559. doi: 10.1371/journal.pone.0008559
- Maddieson, I. (2009). "Calculating phonological complexity," in *Approaches to Phonological Complexity*, eds I. Chitoran, C. Coupé and E. Marsico (Berlin: de Gruyter), 83–110. doi: 10.1515/9783110223958.83
- Matthewson, L. (2014). "The measurement of semantic complexity: how to get by if your language lacks generalized quantifiers," in *Measuring Grammatical Complexity*, eds F. Newmeyer, J. Frederick and L. B. Preston (Oxford: Oxford University Press), 241–263.
- McWhorter, J. (2007). *Language Interrupted: Signs of Non-native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195309805.001.0001
- Meyerhoff, M. (2006). *Introducing Sociolinguistics*. London: Routledge. doi: 10.4324/9780203966709
- Miestamo, M. (2008). "Grammatical complexity in a cross-linguistic perspective," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki, and F. Karlsson (Amsterdam: Benjamins), 23–41. doi: 10.1075/slcs.94.04mie
- Moran, S., and Blasi, D. (2014). "Cross-linguistic comparison of complexity measures in phonological systems," in *Measuring Grammatical Complexity*, eds F. Newmeyer, J. Frederick, and L. B. Preston (Oxford: Oxford University Press), 217–240. doi: 10.1093/acprof:oso/9780199685301.003.0011
- Moran, S., McCloy, D., and Wright, R. (2012). Revisiting population size vs. phoneme inventory size. *Language* 88, 877–893. doi: 10.1353/lan.2012.0087
- Nagy, N., and Meyerhoff, M. (2008). "Introduction: social lives in language," in *Social Lives in language – Sociolinguistics and Multilingual Speech Communities: Celebrating the Work of Gillian Sankoff*, eds M. Meyerhoff, and N. Nagy (Amsterdam: Benjamins), 1–16. doi: 10.1075/impact.24.02nag
- Nettle, D. (2012). Social scale and structural complexity in human language. *Phil. Trans. R. Soc. B* 367, 1829–1836. doi: 10.1098/rstb.2011.0216
- Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226580593.001.0001
- Nichols, J. (2009). "Linguistic complexity: a comprehensive definition and survey," in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 110–125.
- Nichols, J. (2019). "Why is gender so complex? Some typological considerations," in *Grammatical Gender and Linguistic Complexity: Volume I: General Issues and Specific Studies*, eds F. Di Garbo, B. Olsson, and B. Wälchli (Berlin: Language Science Press), 63–92.
- Noglo, K. (2009). "Sociophonetic variation in urban Ewe," in *Variation in Indigenous Minority Languages*, eds J. Stanford, and D. Preston (Amsterdam: Benjamins), 229–244. doi: 10.1075/impact.25.11nog
- Ohala, J. J. (2009). "Languages' sound inventories: the devil in the details," in *Approaches to Phonological Complexity*, eds I. Chitoran, C. Coupé, and E. Marsico (Berlin: de Gruyter), 47–58. doi: 10.1515/9783110223958.47
- Parkvall, M. (2008). "The simplicity of creoles in a cross-linguistic perspective," in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo,

- K. Sinnemäki, and F. Karlsson (Amsterdam: Benjamins), 265–285. doi: 10.1075/slcs.94.17par
- Pellegrino, F., Marsico, E., Chitoran, I., and Coupé, C. (2009). “Introduction,” in *Approaches to Phonological Complexity*, eds I. Chitoran, C. Coupé, and E. Marsico (Berlin: de Gruyter), 1–18. doi: 10.1515/9783110223958.1
- Petré, P., and Anthonissen, L. (2020). Individuality in complex systems: a constructionist approach. *Cogn. Linguist.* 31, 185–212. doi: 10.1515/cog-2019-0033
- Real, F., Chater, N., and Christiansen, M. H. (2018). Simpler grammar, larger vocabulary: how population size affects language. *Proc. R. Soc. B* 285:20172586. doi: 10.1098/rspb.2017.2586
- Rescher, N. (1998). *Complexity. A Philosophical Overview*. New Brunswick; London: Transaction Publishers.
- Sinnemäki, K. (2008). “Complexity trade-offs in core argument marking,” in *Language Complexity: Typology, Contact, Change*, eds M. Miestamo, K. Sinnemäki, and F. Karlsson (Amsterdam: Benjamins), 67–88. doi: 10.1075/slcs.94.06sin
- Sinnemäki, K. (2009). “Complexity in core argument marking and population size,” in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 126–140.
- Szmrecsanyi, B. (2015). “Recontextualizing language complexity.” in *Change of Paradigms—New paradoxes: Recontextualizing Language and Linguistics*, eds J. Daems, E. Zenner, K. Heylen, D. Speelman, and H. Cuyckens (Berlin: De Gruyter), 347–360.
- Szmrecsanyi, B., and Kortmann, B. (2009). Between simplification and complexification: non-standard varieties of English around the world,” in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 64–79.
- Tagliamonte, S. A., and D’Arcy, A. (2009). Peaks beyond phonology: adolescence, incrementation, and language change. *Language* 85, 58–108. doi: 10.1353/lan.0.0084
- Trudgill, P. (2004). Linguistic and social typology: the Austronesian migrations and phoneme inventories. *Linguist. Typol.* 8, 305–320. doi: 10.1515/lity.2004.8.3.305
- Trudgill, P. (2009). “Sociolinguistic typology and complexification,” in *Language Complexity as an Evolving Variable*, eds G. Sampson, D. Gil, and P. Trudgill (Oxford: Oxford University Press), 98–109.
- Trudgill, P. (2011). *Sociolinguistic Typology: Social Determinants of Linguistic Complexity*. Oxford: Oxford University Press.
- Van den Broeck, J. (1977). Class differences in syntactic complexity in the Flemish town of Maaseik. *Lang. Soc.* 6 149–181. doi: 10.1017/S004740450007235
- Zipf, G. K. (1949). *Human Behavior and the Principle of Least Effort*. Boston: Addison-Wesley Press.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Forker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.