

## Layers of morphosyntactic number in Eastern Sudanic

### Project Description

The goal of this project is to investigate phenomena pertaining to the *underexponence* of number (i.e. the presence of fewer than expected number morphemes in a given word) and the *overexponence* of number (i.e. the presence of more than expected number morphemes in a given word) in the severely understudied Eastern Sudanic branch of the Nilo-Saharan language family, and to investigate whether a Cyclic Optimization (CO) approach is better-equipped to account for the attested cross-linguistic variation. The CO model to be developed is one in which syntax builds structure, which is spelled out in a cyclic fashion. At spellout, Optimality Theoretic (OT) constraints determine what type of morphological operations take place, leading to the final exponent (e.g. a given constraint ranking may lead to impoverishment, resulting in underexponence of number). The model combines elements of Kastner (2019), Sande & Jenks (2018) and McPherson & Heath (2016), but is innovative in introducing a separate optimization step for the determination of a phase boundary in the syntax, which then triggers spellout. Furthermore, the CO approach to be developed in this project will be contrasted to two existing CO models for multiple exponence: Caballero & Inkelas' (2013, 2018) approach in Optimal Construction Morphology and Müller's (2020) in Harmonic Serialism, to be further developed in Mor<sup>3</sup>Mor. The bulk of the empirical data for this project will come from original fieldwork conducted in Kenya with speakers of different dialects of four Eastern Sudanic languages (Kalenjin, Turkana, Maa, Didinga), making a significant empirical contribution, since data from these languages have not, to my knowledge, featured in theoretical discussions of number and/or multiple exponence before.

### 1 State of the art and preliminary work

In the last decades there has been significant progress in research on the typology (e.g. Corbett 2000), morphosyntax (e.g. Carstens 1991, Ritter 1991, Wiltschko 2008, Kramer 2016), and semantics (Link 1983, Chierchia 1998, Sauerland 2003, Harbour 2007, de Swart & Farkas 2010) of the grammatical category of number. This research has revealed striking similarities, but also significant differences, in the expression of number distinctions cross-linguistically, both in the nominal and in the verbal domain. Despite this progress, the theoretical implications of the number systems of non-Indo-European languages remain poorly understood. The goal of this research project is to fill in this gap by investigating number phenomena in the severely understudied Eastern Sudanic branch of Nilo-Saharan languages, spoken in East Africa. Eastern Sudanic languages were chosen as the focus of the investigation for two main reasons: a) like other Nilo-Saharan languages, they are well-known for complex systems of number morphology in both the nominal and the verbal domain, (Dimmendaal 2000, 2014, a.o.) and b) they are among the least studied languages of Africa, with even basic descriptions lacking for some of them. Understanding the number systems of these languages is, therefore, of fundamental importance for theoretical treatments of grammatical number in particular, and for theories of the syntax-morphology mapping more generally.

The focus of the investigation will be on phenomena of under- and over-exponence of number morphology, defined as follows:

- (1) **Overexponence:** the existence in a word of multiple exponents spelling out one morphosyntactic feature (in this case, number) and/or syntactic terminal node
- (2) **Underexponence:** the existence in a word of only one morpheme spelling out multiple morphosyntactic features and/or syntactic terminal nodes

These notions are directly related to two broader questions in research on the syntax-morphology interface, namely the treatment of the phenomenon of multiple exponence (defined in (3)) and haplology-like

phenomena, which range from Impoverishment to exponent drop (these phenomena are also discussed in Mor↻Mor in this research unit).

(3) Multiple Exponence (ME) (Harris 2017, 9):

Multiple (or extended) exponence is the occurrence of multiple realizations of a single morphosemantic feature, bundle of features, or derivational category within a word.

In what follows, I first describe the properties of the relevant number phenomena in Eastern Sudanic languages, focusing on the notions of under- and over-exponence. I then discuss previous theoretical approaches to the phenomena in ES languages, before discussing the broader questions: first, how CO has already proved promising for the phenomenon of ME, and second, how previous theories of number morphology in particular point towards CO as the right approach to the relevant phenomena.

*Number in Eastern Sudanic: the empirical landscape*

In the nominal domain, Eastern Sudanic languages exhibit what is called a *tripartite system of number marking* (Dimmendaal 2000, Kouneli 2019), illustrated in (4) with examples from the Surmic language Didinga: some nouns are interpreted as singular in their morphologically unmarked form and form their plural by the addition of a plural suffix, some nouns are interpreted as plural in their unmarked form and form their singular by the addition of a singulative suffix, while a third class of nouns never appear in their unmarked form: they have a singulative suffix in the singular, and a plural suffix in the plural. Kouneli (2020) shows that nouns in all three patterns in (4) show exactly the same syntactic and semantic behavior, with the difference between the three classes lying exclusively in the morphological component. This means that for the plural and singulative marking patterns, we observe underexponence of singular and plural number respectively, in contrast to ‘expected full exponence’ in the replacement pattern. At a descriptive level, we also notice an important difference between Indo-European and Nilo-Saharan languages: while singular number is (almost) always unmarked in the former, it can have an overt exponent in the latter.

(4) Number marking in Didinga (Lohitare et al. 2012)

- |                         |                        |                         |                 |
|-------------------------|------------------------|-------------------------|-----------------|
| a. Plurative marking:   | keem (SG)              | keem- <b>nya</b> (PL)   | ‘camp’          |
| b. Singulative marking: | mining- <b>it</b> (SG) | mining (PL)             | ‘spirit, ghost’ |
| c. Replacement marking: | boloth- <b>it</b> (SG) | boloth- <b>nya</b> (PL) | ‘young bull’    |

While the tripartite system of number marking is limited to nouns in most ES languages, it is also attested in number morphology of adjectives in some languages (e.g. Turkana; Dimmendaal 1983, 2000), while a similar phenomenon is attested in the verbal domain of Nuer (Reid 2019). Reid (2019) claims that Nuer has productive morphology that gives rise to a ‘multiple event’ reading when attached to a verb. This phenomenon is usually called *pluractionality* (cf. Corbett 2000). Even though morphological marking of event plurality is virtually unattested in Indo-European languages, it is relatively common cross-linguistically (see Corbett 2000, Veselinova 2006, Wood 2007 for typological studies, and Henderson 2012, 2017 for a theoretical investigation of the phenomenon in Mayan languages), and widespread in ES languages (Dimmendaal 2014). What is surprising about Nuer is that the pluractional marker has a singulative interpretation when attached to certain verbs. More specifically, certain verbs have plurality encoded in their lexical semantics, with the ‘pluractional’ marker denoting an ‘unexpected number’ reading in this case. This is similar to the situation in the nominal domain: some nouns are morphologically marked in the singular, while others are morphologically marked in the plural.

Previous research on pluractionality has revealed that there is a range of semantic notions conveyed by pluractional morphology cross-linguistically (cf. Corbett 2000, Wood 2007): in addition to the ‘multiple event’ reading, one also finds the conceptually related habitual and intensive readings, while some languages use dedicated morphology to express the notion of multiple participants in an event. What is striking is that some languages use only one morpheme for both multiple events and multiple participants (even when they co-occur semantically), while others require two distinct morphemes. This is illustrated with the difference between Kipsigis and Mupun below: in (5), the plural form of the verb ‘to

hit', *nas*, is used for either multiple events or multiple participants, while when both readings co-occur, no additional morphology is available. In Kipsigis in (6) on the other hand, these notions are all expressed by separate morphology: reduplication expresses the meaning of 'multiple' participants, while the suffix *-to:* is used for multiple participants in an event.

- (5) a. wu **nas** wur  
3.M.SG hit.PST.PL 3SG  
'He hit him many times.'
- b. wu **nas** mo  
3.M.SG hit.PST.PL 3PL  
'He hit them.' OR 'He hit them many times'
- (6) a. **Twal-twa:l-e** là:kwè:t.  
jump.RED-IPFV child.NOM  
'The child is jumping over and over again.'
- b. Twal-**to:s** là:gô:k.  
jump-PL children.NOM  
'The children are jumping.'
- c. **Twal-twal-to:s** là:gô:k.  
jump.RED-PL children.NOM  
'The children are jumping over and over again.'

Thus, while Kipsigis is transparent when it comes to the relationship between morphology and the semantically distinct notions of multiple events and multiple participants in an event, Mupun exhibits underexponence of number: we find one, and not two morphemes, as expected by the semantics.

Pluractionality is not the only area where we find over- and under-exponence of number in the verbal domain of ES languages. Many ES languages display discontinuous agreement when it comes to agreement between the verb and its arguments: number morphology appears in a different agreement slot from person agreement, as shown in (7) for Turkana. While this is not, strictly speaking, overexponence of number (it is rather, overexponence of agreement), in other languages, plural features (in certain cases) do appear in two distinct morphemes, as shown in (8) for participial agreement in Kipsigis.

- (7) ε-raka-si ŋa-atuk ŋa-kaŋ *Turkana* (adapted from Dimmendaal 1985, 135)  
**3-be-PL** cows AGR-mine  
'The cows are mine.'
- (8) a. Mè-á:t ŋò:ktà.  
die-PART dog.NOM  
'The dog is dead.'
- b. Mè-**j(a)-á:t-i:n** ŋò:gì:k  
die-**PL-PART-PL** dogs.NOM  
'The dogs are dead.'

#### *Number in ES: the theoretical landscape*

There is no theoretical investigation of the above described phenomena in ES languages, with the exception of Kouneli (2017, 2019, 2020) on the tripartite system of number marking in Kipsigis and Trommer (2014) on person-number verbal agreement in Turkana. The phenomena, however, have important similarities (and some differences) to better-studied phenomena of great theoretical significance: multiple exponence of number in the nominal domain (e.g. double plurals in languages like Breton and Amharic; Acquaviva 2008, Kramer 2016), plurality in the verbal domain (e.g. pluractionality in Mayan; Henderson 2012, 2017), discontinuous agreement in a number of language families, such as Kartvelian, Algonquian, and Semitic (e.g. Noyer 1992, Halle 2000, Trommer 2001, Harbour 2008, Hewett 2020), and the phenomenon of 'omnivorous number', where only one morpheme is used for number agreement irrespective of whether the verb agrees with one or two plural arguments (e.g.

Nevins 2011, Foley 2017, Lomashvili & Harley 2011, Blix 2020, Bondarenko & Zoppi 2020).

Moving on to previous accounts of these phenomena, a striking observation is that a major theme in all existing work is the division of labor between the syntactic and the morphological component in the explanation of the relevant patterns. In the nominal domain, Kramer's (2016) analysis of double plurals (= number overexponence) and Kouneli's (2019, 2020) analysis of plural haplology (= number underexponence) in Kipsigis both employ a mix of syntactic and post-syntactic morphological operations to account for the data. More specifically, both analyses argue that number features are present on two distinct nodes in the syntax, and post-syntactic operations regulate whether one or both nodes will be spelled out in the morphology. In the verbal domain, whether discontinuous agreement cross-linguistically is a purely syntactic (e.g. Shlonsky 1989), purely morphological (e.g. Noyer 1992), or 'hybrid' phenomenon, where operations in both syntax and morphology conspire to give rise to the final pattern (e.g. Trommer 2001, Harbour 2008) is at the heart of the theoretical debate on the phenomenon, while a mix of syntactic and morphological operations are employed in most analyses of omnivorous number (e.g. Nevins 2011, Lomashvili & Harley 2011, Blix 2020, Bondarenko & Zoppi 2020).

The tension between syntactic and morphological explanations in the literature on over- and underexponence of number indicates that the phenomena to be investigated in this project bear directly on the theory of the syntax-morphology interface; it is clear that both syntactic and morphological operations play an important role and the debate concerns the way in which those interact in the grammar. A major question that the project wants to answer is, thus, the following: what does our theory of the syntax-morphology interface have to look like to explain the cross-linguistic variation in the exponence of number morphology? The goal is to provide an answer to this question by investigating a fresh set of data from a previously unstudied language family. In what follows, I show that the existing literature points towards cyclicity and optimization as key concepts in the theory needed to account for the data, starting with previous CO approaches to multiple exponence.

### *Multiple exponence and Cyclic Optimization*

Recent work has shown that multiple (or extended) exponence (as defined in (3)) is typologically more common than previously thought (Caballero & Harris 2012, Harris 2017). It is, thus, clear that any theory of morphology (and its interfaces) should be able to account for it. In recent work, Caballero & Inkelas (2013, 2018) and Müller (2020) have extensively argued in favor of a CO approach to multiple exponence, even though the details of the two theories have significant differences: Caballero & Inkelas (2013, 2018) develop an analysis in the framework of Optimal Construction Morphology (OCM), which uses a version of Stratal OT, while Müller (2020) develops a Harmonic Serialism account.

Typological work has shown that there are (at least) three types of multiple exponence. In Caballero & Harris' (2012) terms, these are: a) *partially superfluous multiple exponence*: the features spelled out by two exponents stand in a subset relationship, b) *overlapping multiple exponence*: multiple exponents share one feature, but each contributes an additional feature as well, and c) *fully superfluous multiple exponence*: the two exponents spell out exactly the same morphosyntactic features.

Caballero & Inkelas (2013, 2018) predict all three types of multiple exponence to be possible. In their OCM model, words are built bottom-up, step by step, and optimization proceeds from the pressure between faithfulness constraints to a target meaning *M* and morphological well-formedness constraints. There are three major strata (Root - Stem - Word) and the goal is to reach wordhood. Different types of multiple exponence emerge from two core properties of the theory: i) different exponents may have different strengths, meaning that they can satisfy faithfulness constraints to the target meaning *M* to varying degrees (e.g. a synchronically opaque, phonologically weak affix may in a given language not fully satisfy faithfulness constraints to a given target meaning), ii) different exponents satisfy wordhood-related markedness constraints to varying degrees (e.g. a given exponent may only take a root to the stem level, giving rise to the need for an additional exponent to reach word level).

Müller (2020), on the other hand, develops a Harmonic Serialism account of multiple exponence, which predicts that fully superfluous exponence should be unattested (see Mor<sup>↻</sup>Mor for more details on the basics of Müller's 2020 theory). The reason for this is that if an additional exponent cannot dis-

charge any more features than the already present exponents, it could never be the optimal candidate and/or lead to convergence. Partial and overlapping multiple exponence, however, receive a straightforward explanation, since they can be winning candidates resulting from constraints requiring spellout of all relevant morphosyntactic features, which outrank constraints against multiple exponence.

The two theories have the following in common: they both recognize a dedicated morphological component in the grammar, and they both assume that the winning candidate from one cycle serves as the input to the next cycle. One of their most important differences concerns the nature of the cycle: while in OCM there are only three, well-defined strata (root - stem - word), in Harmonic Serialism each change to the input triggers a distinct optimization cycle. Furthermore, the exact nature of the constraints differs between the two theories (e.g. Müller's analysis does not employ any strength-related faithfulness constraints, nor do the notions of root/stem/word play a role in constraint evaluation). As for the relationship between morphology and other components of the grammar, Caballero & Inkelas (2018) explicitly say that they are open to a variety of interface theories, while Müller (2020) assumes that morphology strictly precedes syntax.

### *Number and cyclic optimization*

While CO approaches exist for multiple exponence as a general phenomenon, there are no specific implementations for the type of overexponence of number attested in ES languages. Furthermore, there is no articulated CO theory that can account for the relationship between over- and under-exponence of number of the type investigated in this project. Previous theories of the relevant phenomena, however, do point towards both cyclicity and optimization playing an important role in the correct treatment of the data.

Starting with theories of number morphology in the nominal domain, the functional projection associated with interpretable number in the noun phrase (DP) is usually assumed to be NumP (Ritter 1991 among many others; for the purposes of this discussion, Borer's 2005 DivP could serve the same role, though the theories discussed here make use of nominalizing and verbalizing heads, unlike Borer 2005). It has been argued, however, that number features can be present on other parts of the DP (e.g. Wiltschko 2008), especially on little *n*, i.e. the head responsible for turning roots into nouns frameworks like Distributed Morphology (e.g. Acquaviva 2008, Kramer 2016). In the latter family of approaches, there are several diagnostics used to distinguish between number on Num and number on little *n*:

- number features on Num are interpretable and their interpretation is consistent across languages (e.g. Sauerland et al. 2005, Alexiadou), while number features on little *n* can receive idiosyncratic interpretations (see Acquaviva 2008 for a range of meanings associated with these plurals).
- standard assumptions about locality and roots (e.g. Marantz 1997) predict that the morphological expression and interpretation of number features on little *n* will vary depending on the root, unlike number features on NumP. In other words, NumP plurals are predicted to be regular (both in morphology and semantics), while little *n* plurals are predicted to show a significant degree of irregularity.
- while the primary role of NumP is to host number features, the primary function of little *n* is to nominalize roots; as a result, we expect number features on little *n* to co-exist with (or even take over) other functions (e.g. deriving nominals from verbs or adjectives).

Even though there is evidence for the existence of number features on both Num and little *n* in many languages (e.g. Acquaviva 2008 for a variety of languages, Harbour 2007, 2011 for Kiowa and Jemez, Kramer 2016 for Amharic, Kouneli 2019, 2020 for Nilotic), there is significant cross-linguistic variation in the morphological expression of these two nodes. We have already seen in (4) that in Nilotic languages, number morphology is underexponed in certain contexts. To explain this pattern, Kouneli (2019, 2020) argues that when a little *n* with [ $\alpha$ SG] features co-occurs with a Num node with matching number features, the number node is deleted post-syntactically resulting in exponence of one and not two number morphemes. Amharic is like Nilotic in having plural number features on both little *n* and Num in the syntax (Kramer 2016), but unlike Kipsigis, exponence of both plural nodes is

possible (e.g. we find the form *k'al-at-otftf* for the plural noun 'words' in Amharic, where *-at* spells out a plural little *n* and *-otftf* spells out a plural Num node), meaning that the haplology operation proposed by Kouneli (2019, 2020) does not apply. Kouneli's analysis indicates that there is a constraint against multiple exponence operating in Nilotic languages. Variants of such a constraint appear in many previous theories that use optimization, irrespective of whether they additionally use cyclicity or not. Specific examples include *Simple Exponence* ("prevent Vocabulary Insertion of a vocabulary item that would lead to multiple exponence of any feature") in Foley (2017) or *Minimize Exponence* ("the most economical derivation will be the one that maximally realizes all the formal features of the derivation with the fewest morphemes") in Siddiqi (2006, 14, 162). Thus, the cross-linguistic variation in the morphology of number suggests an optimization approach: while both Num and little *n* could carry number features in the syntax, the relative ranking of constraints against multiple exponence, on the one hand, and constraints favoring the spellout of all syntactic features, on the other hand, can be used to explain different patterns of morphological exponence.

There is evidence that cyclicity also plays a significant role in the morphosyntax of number. For example, as was mentioned earlier, it has been observed that little *n* plurals may be morphologically irregular and subject to idiosyncratic interpretations, while Num plurals are regular morphologically and semantically (e.g. Acquaviva 2008, Kramer 2016). The notion of cyclicity has been prominent in accounts of this observation: it has been claimed that *n* introduces a phase boundary, making any effect of the root on morphophonology and interpretation opaque to elements outside of that phase (e.g. Marantz 2001, Embick 2010; see also Fenger & Harđarson 2018 for an application of this idea to number).

Furthermore, for those languages with robust evidence for two locations for number features, there are cross-linguistic differences in the accessibility of number on *n* to operations in later cycles. In both Kipsigis (Nilotic) and Kiowa, for example, there are number features on both *n* and Num (Harbour 2007, Kouneli 2019). However, we find a difference in agreement relations in the two languages: in Kipsigis, D only agrees with number features on Num (i.e., the highest node with number features) (Kouneli 2019, 2020), while in Kiowa, D agrees with number features on both Num and little *n*, resulting in the presence of an 'inverse' marker when the features of the two nodes bear opposite values (e.g. +SG and –SG) (Harbour 2011). This is illustrated in the examples below. There is independent evidence (not shown here) that both the Kipsigis noun in (9) and the Kiowa noun in (10) have [+SG] features on little *n* (Harbour 2011, Kouneli 2020). In both examples below, the semantic interpretation of the noun is plural, indicating the presence of a [–SG] Num node above the [+SG] little *n* node. We see in (9) that the adjective displays plural agreement in Kipsigis, which means that it agrees with the [–SG] features on Num, with the number features on little *n* being inaccessible to agreement operations. In Kiowa, on the other hand, we see in (10) that the same syntactic configuration gives rise to an inverse marker, which in Kiowa indicates agreement with features bearing opposite values (i.e. simultaneous agreement with a [+SG] and a [–SG] feature); this means that the number features on little *n* are accessible to later cycles in Kiowa, unlike Kipsigis.

- (9) Tórô:r-è:n là:gô:k.  
 tall-PL girls.NOM  
 'The girls are tall.'

*Kipsigis* (Kouneli 2020: 35)

- (10) óópíí-dó e-dóó  
 fish-INV 3INV-be  
 'It's some fish.'

*Kiowa* (Harbour 2011: 564)

Moving on to the second phenomenon under investigation in this project, pluractionality (or 'verbal number'), theoretical research on the phenomenon is limited. Previous theoretical work is mostly concerned with the semantics of pluractionality (e.g. Henderson 2012), while some studies investigate locality conditions in the determination of verbal suppletion and reduplication, two common morphological reflexes of pluractionality (e.g. Bobaljik & Harley 2017). In these studies, however, there has been no attempt to provide a systematic explanation of how the semantics, syntax, and morphology interact in the phenomenon of pluractionality. There are two exceptions: Thornton (2019) and Amato (2018)

give a preliminary account of the relationship between the syntactic representation of pluractionality and its morphological expression. These studies, however, make two assumptions about the syntax of verbal number that are shown to be inadequate for the data from the ES languages under investigation in this project. Both Thornton (2019) and Amato (2018) assume that there is only one syntactic head associated with pluractional morphology. Thornton (2019) uses this to explain why suppletion and reduplication are in complementary distribution in the languages she investigates, while Amato (2018) uses one head to explain why only one morpheme is used for the two notions of ‘multiple events’ and ‘multiple participants in an event’, as shown in (5) earlier. Both of these empirical assumptions, however, turn out to be wrong for a number of ES languages. First, it was already shown in (6) that in Kipsigis, we do indeed find two distinct morphological realizations of the two semantic notions involved in pluractionality, while (11) shows that suppletion and reduplication are not in complementary distribution in Kipsigis: the root for ‘run’ is suppletive for plural subjects, but it can be reduplicated to indicate the notion of multiple events.

- (11) a. labat-i            là:kwè:t.  
run.SG-IPFV child.NOM  
‘The child is running.’  
b. Ruaj            là:gô:k.  
run.IPFV.PL children.NOM  
‘The children are running.’  
c. Ruaj-ruaj            là:gô:k.  
run.IPFV.PL.RED children.NOM  
‘The children are running over and over again.’

The data above show that verbal number presents a similar picture to the tension between over- and under-exponence of number morphology in the nominal domain, which is in line to the semantic parallelism between nominal and event plurality that we observe cross-linguistically (Henderson 2012). An approach in terms of optimization is, thus, likely to be successful in this case as well, with different rankings of faithfulness constraints and markedness constraints against multiple exponence giving rise to the type of contrast that we see between Kipsigis-like and Mupun-like languages. Given the paucity of research on pluractionality, it is difficult to find previous accounts that crucially rely on cyclicity, but there is some preliminary evidence for its role. First, in Henderson’s (2012) analysis of pluractionality in Mayan languages, it is important that there are two different locations for plural morphology, with one being closer to the root. Second, in Bobaljik & Harley (2017), locality plays an important role in the determination of verbal number morphology. Furthermore, in those cases where verbal number morphology is related to multiple participants, it is not always easy to distinguish it from regular agreement for number features. For example, there is evidence from Hiaki that deep argument structure relations matter for verbal number, while surface relations after movement matter for agreement (Bobaljik & Harley 2017). What is meant though by ‘deep’ versus ‘surface’ and how do we account for cross-linguistic variation? These observations in the literature make a connection to cycles possible.

The connection between verbal number and agreement takes us to the last domain of investigation of the project: multiple layers of number in agreement between the verb and its arguments. Two distinct, but possibly related, phenomena are relevant: multiple exponence of number morphology in agreement and omnivorous number, where a plural morpheme is used irrespective of which argument (subject vs. object) is plural; when both arguments are plural, we still see one and not two morphemes. These phenomena have received a lot of attention in the theoretical literature. Multiple exponence has been the subject of two different CO theories: a version of Stratal OT in Caballero & Inkelas (2013) and Harmonic Serialism in Müller (2020), as discussed in the previous section. As for omnivorous number (which has mostly been investigated in Kartvelian languages), cyclicity has been prominent in previous theoretical approaches in two respects: a) by using a cyclic implementation of the operation of Agree to capture agreement with multiple arguments (Béjar & Řezáč 2009, Nevins 2011), and b) in various implementations of cyclic spellout in theories that interleave syntax and morphology (Blix 2020, Bondarenko & Zompì 2020). Optimization has been less prominent in approaches to omnivorous number,

but, as has already been mentioned, Foley (2017) uses OT constraints to account for the phenomenon in Georgian, while Müller (2009) is an example of an optimization approach to agreement.

The empirical landscape in the domain of number agreement, as was already discussed, shows the same tension between under- and over-exponence of number morphology that we find elsewhere, and suggests that both cyclicity and optimization play an important role in determining the observed cross-linguistic variation. In combination to the range of previous theoretical approaches that use cyclicity and/or optimization discussed in this section, a cyclic optimization approach suggests itself for the explanation of the relevant phenomena.

## 1.1 Project-related publications

- Kouneli, M. 2017. The morphological expression of number in Kalenjin. In A. Lamont & K. Tetzloff (eds.), *Proceedings of the 47th Annual Meeting of the North East Linguistic Society (NELS)*, Vol.2: 167-176. Amherst: GLSA.
- Kouneli, M. 2018. Plural marking on mass nouns: Evidence from Greek. In M. Dali, E. Mathieu & G. Zareikar (eds.), *Gender and noun classification*. Oxford: Oxford University Press.
- Kouneli, M. 2019a. *The syntax of number and modification: An investigation of the Kipsigis DP*. Ph.D. Dissertation, New York University.
- Kouneli, M. 2019b. Determiner spreading and modification in Kipsigis. In *Proceedings of the 54th Meeting of the Chicago Linguistic Society (CLS 54)*.
- Kouneli, M. 2020. Number-based noun classification: the view from Kipsigis. *Natural Language & Linguistic Theory*.

## 2 Objectives and work programme

### 2.1 Objectives

The goal of this project is twofold: first, it aims at making a significant empirical contribution by providing a detailed investigation of number (both in the nominal and in the verbal domain) in ES languages, thus bringing fresh data from an understudied language family to theoretical debates about grammatical number. Second, it aims at developing a CO approach to number phenomena that goes beyond the narrow scope of number and extends to what the architecture of the syntax-morphology interface looks like. This latter goal has a comparative angle, as the CO model to be developed in this project contrasts with the theories of the syntax-morphology interface explored in Mor<sup>↔</sup>Mor and Sem<sup>↔</sup>Phon. Overall then, the goals of the project are part of the broader objectives of this RU: extend the empirical coverage of CO theories and compare the formal properties of different CO models.

The view of the syntax-morphology interface to be developed in this project shares a number of assumptions with Distributed Morphology (Halle & Marantz 1993). More specifically, the grammar follows the Y-model and syntax is the only structure-building module, which operates on phonology-free morphosyntactic features/feature bundles. I assume that morphological words correspond to complex heads (Halle & Marantz 1993, Arregi & Nevins 2012, a.o.); in other words, Head Movement creates words. Once all syntactic operations are complete, certain morphological operations (e.g. Impoverishment) can manipulate feature bundles. However, in this model, no operations that mimic movement are allowed in the post-syntactic component (e.g. Local Dislocation or Lowering), with all movement being confined to the syntax.

After these operations have taken place, vocabulary items are inserted in each terminal node. In standard DM, optimization plays a role in Vocabulary Insertion, which follows the Subset Principle (Halle 1997): the phonological exponent of a vocabulary item can be inserted if the item contains all or a subset of the features present at the terminal node, while at the same time the item must have no feature that is absent from the node. In other words, vocabulary items compete for insertion and the most specific one wins. While other types of optimization (e.g. OT constraint evaluation) and cycles are not an inherent property of DM, a number of implementations of the theory assume one or the other (or both). DM coupled with phase theory (e.g. Embick 2010) crucially relies on the



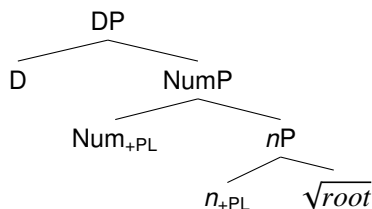
notion of cyclicity, while there are various Optimality Theory (OT) implementations of DM (e.g. Trommer 2001, Rolle 2019); phases coupled with an OT implementation give rise to CO models (e.g. Kastner 2019). The specific model to be developed in this project works as follows: OT constraints at each phrasal node determines whether said node introduces a phase boundary; if yes, the structure is sent to spellout, where optimization determines what type of morphological operations (e.g. Impoverishment) take place. In the final step, Vocabulary Insertion takes place, which follows the Subset Principle. Furthermore, Vocabulary Insertion is assumed to take place ‘inside-out’, i.e. starting from the most embedded node (Bobaljik 2012). The model combines elements of Keine & Müller (2015), Kastner (2019), Sande & Jenks (2018), and McPherson & Heath (2016), but is innovative in introducing a separate optimization step for the determination of a phase boundary. This latter step is a particular formalization of the idea behind Dynamic Phases (Bobaljik & Wurmbrand 2005, 2013). The following schema summarizes the essential steps of the model:

- (12) Syntax (Merge-based and feature-driven) → Optimization Cycle 1 (OT constraints): phase boundary → Optimization Cycle 2 (OT constraints): post-syntactic operations → Optimization Cycle 3 (Subset Principle): Vocabulary Insertion

Coming back to the phenomena at hand, the main hypothesis entertained is the following: in certain configurations (which could vary by language), number features are present in two different nodes in the syntax. Spelling out both nodes gives rise to overexponence, while spelling out only one node gives rise to underexponence. Which node(s) will be spelled out depends on different types of constraint interaction post-syntactically. Relevant constraints are faithfulness constraints requiring spellout of all morphosyntactic features and markedness constraints against multiple exponence; however, there are certainly more constraints involved, such as locality considerations (e.g. deletion may be more likely for adjacent nodes) and the exact featural make-up of the nodes involved (e.g. deletion may be less likely if one of the nodes hosts other features in addition to number).

I illustrate here how the system would work for nominal number, which is the area for which we have a better understanding. I will assume the simple structure in (13):

- (13) DP with number features on both *n* and Num



The system works in the following steps:

1. All languages have number features on Num; languages vary in whether they have number features on little *n* (this is a lexical difference).
2. Every time Merge takes place, an optimization step applies where it is determined whether the resulting phrase is a phase (the relevant constraints are related at least to the features on the head of the phrase, but further factors will be explored); I assume that little *n* (and little *v* in the verbal domain) are always phases (Marantz 1997 and subsequent work) possibly due to unviolated constraints related to root categorization.
3. If X is a phase head, spellout is triggered, which happens in two steps.
4. In the first step, morphological operations (e.g. Impoverishment) take place; these operations emerge as results of constraint rankings.
5. In the second step, Vocabulary Insertion takes place.

Cross-linguistic variation comes from differences in the steps above: syntax (how many number nodes are present and what their role is), phasehood (which nodes introduce phase boundaries), morphological operations (which morphological operations take place), and Vocabulary Insertion (what

vocabulary items are available). Optimization plays a role in Steps 2 and 4 above, where different constraint rankings will give rise to different exponence patterns, and with the Subset Principle in Vocabulary Insertion. For example, the difference between Kipsigis (where we find underexponence of number) and Amharic (where we find overexponence) could be due to different rankings of a faithfulness constraint (favoring spellout of all morphosyntactic features) and a markedness constraint against multiple exponence; these constraints would operate after spellout and before vocabulary insertion. Similarly, the difference between Kipsigis and Kiowa agreement could be due to a difference between the two languages in the status of the Num head: it might be a phase head in Kipsigis (thus, rendering little *n* inaccessible to later cycles), but not in Kiowa. Which constraints determine the phase status of a head will be one of the major topics of investigation of this project, but possible candidates include the particular featural make-up of a head and whether this head is crucial for the interpretation of another head (cf. Bobaljik & Wurmbrand 2013).

The same model can be extended to the less studied phenomenon of pluractionality, assuming a parallelism between nominal and verbal number in terms of individual vs. event plurality (Henderson 2012). Before developing the theoretical model, a detailed investigation of the properties of the phenomenon in ES languages will be carried out. Assuming a noun - verb parallelism, we can extend the diagnostics for number features on little *n*, discussed previously, to number features on little *v*, the equivalent lexical head in the verbal domain.

This leaves us with verbal agreement for number, which is the most complex of the phenomena to be investigated in this project; this is because we are not dealing with a small number of projections close to the root (as in nominal number and pluractionality), but with multiple projections in the clausal spine. Furthermore, agreement in the verbal domain is affected by a number of additional syntactic operations, such as movement and case assignment. For example, all of the ES languages under investigation in this project are verb-initial, which means that the verb (either as a head or as part of a VP) almost certainly moves to a higher projection in the clause. The broad mechanisms of the CO model, however, are the same in this case as well: the central assumption is that in ES languages, agreement between the verb and its arguments is realized in two distinct heads in the syntax, and it is different rankings of phase-related and morphology-related constraints that give rise to different patterns of morphological exponence of number. There are cross-linguistic differences in whether a language has agreement in two places in the syntax or in one, similarly to how some languages have number features on both Num and little *n* (e.g. Kipsigis), while others only have them on Num (e.g. English).

This particular CO model makes a number of predictions about the phenomena under investigation; the most important hypotheses are summarized below:

- **H1:** The presence of multiple number nodes in the syntax predicts that multiple exponence of number should correlate with syntactic effects, i.e. it is not exclusively a morphological phenomenon. Possible syntactic effects include interaction with movement (e.g. the presence of a number node could be an intervener for Head Movement), interaction with agreement (e.g. an additional node could act as either a probe or a goal for agreement, depending on the configuration), and possibly binding relations (if the node hosting number features can, for example, attract nominals to its specifier, affecting the surface configuration between DPs).
- **H2:** Since hapology operations can take place post-syntactically in this model, we predict a split in the languages that display only one number morpheme: there should be languages where we find the same syntactic effects as in the languages with multiple exponence (=languages with two number morphemes but post-syntactic Impoverishment) and languages where “underexponence” simply reflects the presence of only one node in the syntax.
- **H3:** Multiple Exponence is explained in this model as the presence of different nodes in the syntax, which necessarily have different functions. Thus, *fully superfluous exponence* in the terminology of Caballero & Harris (2012) is predicted to be impossible in my theory. This is a shared prediction with the Mor<sup>☞</sup>Mor project, though the explanation for it is different.
- **H4:** The Harmonic Serialism model in Mor<sup>☞</sup>Mor predicts that in cases of *partially superfluous* Multiple Exponence (i.e. the features of the two exponents stand in a subset relation), the less specific exponent always appears closer to the root. No such prediction is made in my theory,

which allows both this order and the reverse, where the less specific exponent is further from the root.

- **H5:** In this project, movement is restricted to the syntactic component, and cannot apply post-syntactically. It is, thus, predicted that the relative ordering of the number (and other) morphemes will be determined by syntactic factors. This is different from the claim in Mor<sup>3</sup>Mor, where purely morphological movement is possible.

Finally, the theory to be developed in this project predicts certain theory-internal correlations. For example, if the two relevant number nodes are separated by a phase boundary, morphological operations applying to them would be determined in distinct optimization cycles, and haplogy would no longer be possible. Thus, it is predicted that number nodes belonging to different phases will always lead to overexponence.

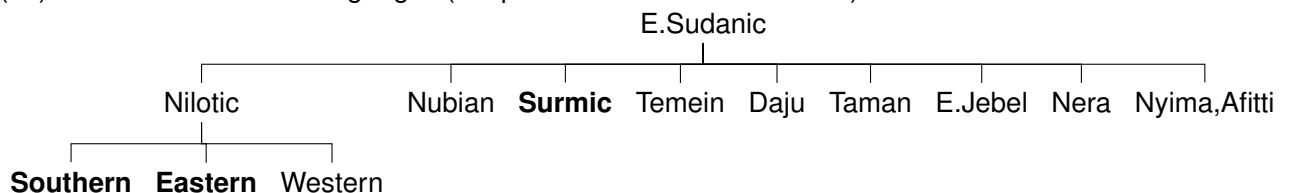
## 2.2 Work programme incl. proposed research methods

As has already been discussed, the bulk of the empirical data considered in this project will come from fieldwork on different dialects of four ES languages: Kalenjin (Southern Nilotic), Maa (Eastern Nilotic), Turkana (Eastern Nilotic), and Didinga (Surmic). A detailed investigation of the relevant phenomena in these languages will be carried out, along with a theoretical analysis couched in the CO model outlined in the previous section. Then, the theory will be tested against the typological and theoretical generalizations already available in the literature, with the hope of providing a unified theory. In this section, I give details on the languages to be investigated and the reason why they were chosen, I discuss fieldwork methodology, and I give a tentative timeline.

### 2.2.1 Languages of investigation

Given the paucity of descriptive materials in many ES languages, the languages chosen as the focus of the investigation in this project are those for which basic grammatical descriptions are already available and report some (or all) of the relevant number-related phenomena. Furthermore, languages were chosen for which speakers can be found in a single country (Kenya), thus making the data collection process easier. More details for each language are given below, while (14) shows the internal classification of ES languages; fieldwork will be conducted on languages from the Southern and Eastern Nilotic and Surmic branches, which are bolded.

(14) Eastern Sudanic languages (adapted from Dimmendaal 2000)



- **Kalenjin dialects (Kipsigis, Nandi, Tugen, Keiyo, Endo-Marakwet):** Kalenjin dialects belong to the Southern Nilotic branch of ES and are spoken by approximately 5 million speakers in Kenya (Eberhard et al. 2020). Descriptions of some dialects are available (e.g. Toweett 1979 for Kipsigis, Creider & Creider 1989 for Nandi, Zwarts 2004 for Endo-Marakwet), but they are understudied in the theoretical literature. These dialects have been reported to display all phenomena under investigation in this project (multiple layers of number in the nominal domain, pluractionality both for multiple events and multiple participants, multiple exponence of number in verbal agreement, and discontinuous agreement for person/number), and micro-comparative work in the dialects is a good testing ground for our theory.
- **Samburu and Maasai dialects of Maa:** The Maa dialects belong to the Eastern Nilotic branch of ES and are spoken by approximately 2 million speakers in Kenya (Eberhard et al. 2020). Samburu and Maasai are the two dialects with most speakers, and will be the ones studied in

this project. They are relatively well-described (e.g. Tucker & Mpaayei 1955), and there are some theoretical studies (e.g. Koopman 2003). It displays all of the relevant phenomena: the tripartite system of number marking in the nominal domain, pluractionality in the verbal domain, as well as discontinuous exponence of person and number agreement in the verbal domain.

- **Turkana language:** Turkana belongs to the Eastern Nilotic branch of ES and is spoken by approximately 1 million speakers in Kenya (Eberhard et al. 2020). It is relatively well-described (e.g. Dimmendaal 1983), but it is understudied in the theoretical literature, and there is, thus, a lack of syntactic data. It displays all of the relevant phenomena: the tripartite system of number marking in the nominal domain, pluractionality in the verbal domain, as well as discontinuous exponence of person and number agreement in the verbal domain (Dimmendaal 1983). Furthermore, it displays an inverse system of person marking, which indicates that the verb agrees with multiple arguments (Trommer 2014).
- **Didinga language:** Didinga belongs to the Surmic branch of ES, and is thus the only non-Nilotic language in the sample. It is primarily spoken in South Sudan by approximately 60,000 people (Eberhard et al. 2020), but there is a significant community of Didinga-speaking refugees living in Nairobi. Lohitare et al. (2012) report the Nilo-Saharan tripartite system of number marking in the nominal domain for Didinga, as well as a pattern of discontinuous exponence of person and number morphology in subject-verb agreement. It is unclear if Didinga has pluractionality. There are descriptive materials on Didinga (Lohitare et al. 2012), but there is a poor understanding of syntactic processes.

Even though the focus will be on these languages, existing descriptions of other ES languages will be consulted, and it is the goal of the project to provide a theory that can account for variation of the phenomena in the family.

## 2.2.2 Fieldwork methodology

Fieldwork on the languages described above will be conducted in Kenya, with Nairobi (the capital city) as the base of the investigators. The goal is to work with at least 5 speakers from each language/dialect and to collect data on the number phenomena under investigation. There are no significant obstacles to locating speakers who will be willing to participate in the study, and I have conducted fieldwork in Kenya before. Sessions with the linguistic consultants will only begin after informed consent is obtained: the consultants will be informed by the investigator(s) on the details of the study, and permission will be asked for participation and/or audio recording; (non-)anonymity will also be discussed. Participants will have the option of withdrawing from the study at any time without cost. Consultants will be compensated on an hourly basis for their participation.

Data collection will primarily consist of targeted elicitation interviews (with English, one of the official languages of Kenya, as the metalanguage), but text collection and observation of simultaneous conversations (whenever possible) will also be conducted. Elicitation interviews will consist of planned stimuli aiming at understanding the morphology, syntax, and semantics of the relevant number phenomena in a given language. Sample questions that the investigator will be asking the linguistic consultants are the following:

- How would you say [insert English sentence] in your language?
- Is this sentence [insert sentence in investigation language] a natural sentence in your language? If not, how would you change it?
- Imagine the following conversation scenario; would [insert sentence in investigation language] be an acceptable utterance in this context?

If the consultants agree to it, all sessions will be recorded (audio only). It is also possible that some interviews will be conducted online (while the investigators are in Germany), but such interviews are not ideal due to difficulties in the correct transcription of the complex morphophonology of ES languages (especially tone).

### 2.2.3 Timeline

The PhD student and the post-doctoral researcher in this project will be involved in all research stages of the project (including fieldwork), where each investigator will be in charge of one language/dialect cluster. The following is a tentative timeline.

#### Year 1

- Careful literature review of the relevant phenomena (both on the theoretical and the empirical side): number morphology in the nominal domain (with emphasis on double plurals in ES languages and cross-linguistically), pluractionality, discontinuous agreement, omnivorous number, as well as grammatical descriptions of the four ES languages under investigation
- Preparation for fieldwork:
  - collection of all available materials on the four languages under investigation and review of the relevant theoretical literature
  - design of stimuli (for elicitation interviews) to be collected in Kenya
  - practical matters (e.g., affiliation with Kenyan university and application for research permit, initiate contact with the native speakers who will act as consultants, travel plans etc.)
- First field trip to Kenya at the end of the year: collect data on grammar of number in all languages under investigation

#### Year 2

- Processing and analysis of the data collected in Kenya
- Logging of fieldnotes and recordings to DOBES database (see Section 2.4 for more details on the database)
- Development of theoretical analysis for collected data
- Preparation for second fieldwork trip (to collect data needed to test the predictions of the theory as well as any data points that we missed in the first field trip)
- Second fieldwork trip to Kenya

Year 3

- Processing and analysis of the data collected in the second fieldtrip to Kenya
- Entry of new data to DOBES database
- Extension of the theory to phenomena outside of ES languages and refinement of the CO approach

Year 4

- Further refinement of the theory
- Publication of the DOBES database
- Presentation of the results in conferences and publications

### 2.3 Other information – Cooperation within the RU

**Mor<sup>Phon</sup>**: One of the empirical domains of the Mor<sup>Phon</sup> project is the morphological tone system of Nilotic languages, and there will be collaboration in the data collection process; the post-doctoral researcher of the Mor<sup>Phon</sup> project will participate in one of the two planned fieldtrips to Kenya (more details are given in later sections). On the theoretical side, we will work together on modeling patterns of ‘underexponence’, which are manifested in tonal overwriting patterns in the Morph<sup>Phon</sup> project and number morphology in this project.

**Mor<sup>Mor</sup>**: The phenomena investigated in this project are relevant case studies for the sub-topics of Impoverishment, Exponent Drop, and Morphological Movement in the Mor<sup>Mor</sup> project, since they constitute examples of multiple exponence, discontinuous exponence, and exponent drop. The two projects develop different CO approaches to these phenomena, and theoretical comparison will be made (different predictions made by the two theories were already discussed in Section 2.2). Two theoretical differences are prominent: first, syntax strictly precedes morphology in my model, while the reverse is true in Mor<sup>Mor</sup>; second, syntax is the only structure-building component in my theory, while both morphology and syntax are structure-building components in Mor<sup>Mor</sup>.

**Syn<sup>Phon</sup>**: One of the empirical domains of the Syn<sup>Phon</sup> project is the cross-linguistic variation in the behavior of monosyndetic and polysyndetic coordination. The appearance of one vs. many overt coordinators bears striking similarities (at least at a descriptive level) to the notions of under- and over-exponence under investigation in this project. Furthermore, theories exist that postulate multiple heads in the syntax to explain coordination patterns (see Syn<sup>Phon</sup> for details). Given this theoretical background and the shared assumption with the Syn<sup>Phon</sup> project that there is a post-syntactic component that can manipulate feature bundles, it seems possible to develop a unified theory of under- and over-exponence in the two empirical domains, number and coordination.

**Syn<sup>Syn</sup>**: The Syn<sup>Syn</sup> project will develop a CO theory of syntactic repairs. Some of the phenomena investigated in this project can be conceptualized as repairs, especially underexponence of number in the domain of agreement, as is also discussed in the Syn<sup>Syn</sup> project. Underexponence in number agreement could potentially provide an opacity-based argument for a CO approach to repairs; provided that the constraints favoring deletion of a node/feature depend on notions like locality, syntactic operations could (in the right configuration) counterbleed post-syntactic deletion. Furthermore, in investigating verbal agreement in Eastern Nilotic languages in particular (Turkana and the Maa dialects), the analysis of number agreement will necessarily also depend on the analysis of the inverse system of these languages; the verb agrees with both the subject and the object, and an ‘inverse’ marker is inserted in subject-object configurations that violate certain person prominence hierarchies. Inverse markers are well-known repairs, but have been virtually unstudied in Nilotic languages. The detailed investigation of verbal agreement in these languages that will be carried out in this project will provide useful background for the classification and analysis of this particular repair phenomenon in the Syn<sup>Syn</sup> project.

### 3 Bibliography

- Acquaviva, P. 2008. *Lexical plurals: A morphosemantic approach*, vol. 19. Oxford: Oxford University Press.
- Alexiadou, A. 2000. Morphological and semantic markedness revisited: The realization of plurality across languages. *Zeitschrift für Sprachwissenschaft* 38. 123–154.
- Amato, I. 2018. A morphosyntactic account of verbal number in Mupun. In *Proceedings of the 26th Conference of the Student Organization of Linguistics in Europe*, 20–43. Leiden: Leiden Centre for Linguistics.
- Arregi, K. & A. Nevins. 2012. *Morphotactics: Basque auxiliaries and the structure of spellout*. Heidelberg: Springer.
- Béjar, S. & M. Řezáč. 2009. Cyclic Agree. *LI* 40. 35–73.
- Blix, H. 2020. Spans in South Caucasian agreement: Revisiting the pieces of inflection. *NLLT* 1–55.
- Bobaljik, J. 2012. *Universals in comparative morphology. suppletion, superlatives, and the structure of words*. Cambridge, MA: MIT Press.
- Bobaljik, J. & H. Harley. 2017. Suppletion is local: Evidence from Hiaki. In H. Newell, M. Noonan, G. Piggott & L. Travis (eds.), *The structure of words at the interfaces*, 141–159. Oxford: Oxford University Press.
- Bobaljik, J. D. & S. Wurmbrand. 2005. The domain of agreement. *Natural Language & Linguistic Theory* 23. 809–865.
- Bobaljik, J. D. & S. Wurmbrand. 2013. Suspension across domains. In A. Marantz & O. Matushansky (eds.), *Distributed Morphology today: Morphemes for Morris Halle*, 185–198. Cambridge MA: MIT Press.
- Bondarenko, T. & S. Zoppi. 2020. Leftover Agreement: Spelling out Kartvelian number. Handout of talk given online at GLOW 2020.
- Borer, H. 2005. *Structuring sense: In name only*, vol. 1. Oxford: Oxford University Press.
- Caballero, G. & A. Harris. 2012. A working typology of multiple exponence. In F. Kiefer et al. (eds.), *Current issues in morphological theory: (Ir)regularity, analogy, and frequency*, 163–188. Amsterdam: John Benjamins.
- Caballero, G. & S. Inkelas. 2013. Word construction: Tracing an optimal path through the lexicon. *Morphology* (23). 103–143.
- Caballero, G. & S. Inkelas. 2018. A construction-based approach to multiple exponence. In G. Booij (ed.), *The construction of words. Advances in Construction Morphology*, 111–139. Dordrecht: Springer.
- Carstens, V. 1991. *The morphology and syntax of determiner phrases in Kiswahili*. Los Angeles: UCLA dissertation.
- Chierchia, G. 1998. Reference to kinds across language. *Natural language semantics* 6(4). 339–405.
- Corbett, G. G. 2000. *Number*. Cambridge: Cambridge University Press.
- Creider, C. & J. T. Creider. 1989. *A grammar of Nandi*. Hamburg: Helmut Buske Verlag.
- Dimmendaal, G. J. 1983. *The Turkana language*. Dordrecht: Foris.
- Dimmendaal, G. J. 1985. Prominence hierarchies and Turkana syntax. In G. J. Dimmendaal (ed.), *Current approaches to African linguistics*, 127–148. Dordrecht: Foris.
- Dimmendaal, G. J. 2000. Number marking and noun categorization in Nilo-Saharan languages. *Anthropological linguistics* 214–261.
- Dimmendaal, G. J. 2014. Pluractionality and the distribution of number marking. In G. J. Dimmendaal (ed.), *Number: Constructions and semantics. Case studies from Africa, Amazonia, India and Oceania*, 57–75. Amsterdam: John Benjamins.
- Eberhard, D. M., G. F. Simons & C. D. Fennig (eds.). 2020. *Ethnologue: Languages of the world. Twenty-third edition*. Dallas: SIL International.
- Embick, D. 2010. *Localism versus globalism in morphology and phonology*. Cambridge, MA: MIT Press.
- Fenger, P. & G. R. Harðarson. 2018. Your ns are numbered! On linking morphemes in Dutch. In *Proceedings of the Linguistic Society of America*, vol. 3, Article 26, 1–15.
- Foley, S. 2017. Morphological conspiracies in Georgian and optimal vocabulary insertion. In *Proceedings of CLS 52*, Chicago: Chicago Linguistic Society.
- Halle, M. 1997. Distributed Morphology: Impoverishment and Fission. In B. Bruening, Y. Kang & M. McGinnis (eds.), *MIT working papers in linguistics 30: Papers at the interface*, 425–449. Cambridge, MA: MITWPL.

- Halle, M. 2000. Distributed Morphology: Impoverishment and fission. In J. Lecarme, J. Lowenstamm & U. Shlonsky (eds.), *Research in afroasiatic grammar: Papers from the 3rd conference on Afroasiatic languages in Sophia Antipolis, France, 1996*, 125–149. Amsterdam & Philadelphia: John Benjamins.
- Halle, M. & A. Marantz. 1993. Distributed Morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.), *The view from building 20*, 111–176. Cambridge MA: MIT Press.
- Harbour, D. 2007. *Morphosemantic number: From Kiowa noun classes to UG number features*, vol. 69. Springer Science & Business Media.
- Harbour, D. 2008. Discontinuous agreement and the syntax-morphology interface. In D. Harbour, S. Béjar & D. Adger (eds.), *Phi-theory: Phi-features across modules and interfaces*, 185–220. Oxford: Oxford University Press.
- Harbour, D. 2011. Valence and atomic number. *LI* 42(4). 561–594.
- Harris, A. 2017. *Multiple Exponence*. Oxford: Oxford University Press.
- Henderson, R. 2012. *Ways of pluralizing events*. Santa Cruz: UC Santa Cruz dissertation.
- Henderson, R. 2017. Mayan pluractionality. In J. Aissen, N. England & R. Zavala (eds.), *The Mayan languages*, Oxford: Oxford University Press.
- Hewett, M. 2020. Discontinuous agreement in Semitic: A fission-based approach. Manuscript, University of Chicago.
- Kastner, I. 2019. Templatic morphology as an emergent property. *NLLT* 37(2). 571–619.
- Keine, S. & G. Müller. 2015. Differential argument encoding by impoverishment. In I. Bornkessel-Schlesewsky, A. Malchukov & M. Richards (eds.), *Scales and hierarchies: A cross-disciplinary perspective*, 75–130. Berlin: de Gruyter.
- Koopman, H. 2003. The locality of agreement and the structure of the DP in Maasai. In W. E. Griffin (ed.), *The role of agreement in natural language: TLS 5 proceedings*, vol. 53, 207–227. Texas Linguistics Forum.
- Kouneli, M. 2017. The morphological expression of number in Kalenjin. In A. Lamont & K. Tetzloff (eds.), *Proceedings of the 47th Annual Meeting of the North East Linguistic Society (NELS)*, vol. 2, 167–176. Amherst: GLSA.
- Kouneli, M. 2019. *The syntax of number and modification: An investigation of the Kipsigis DP*. New York: New York University Ph.D. Dissertation.
- Kouneli, M. 2020. Number-based noun classification: The view from Kipsigis. *Natural Language & Linguistic Theory*.
- Kramer, R. 2016. A Split Analysis of Plurality: Number in Amharic. *LI* 47(3). 527–559.
- Link, G. 1983. The logical analysis of plurals and mass terms: A lattice-theoretic approach. In R. Bäuerle, C. Schwarze & A. von Stechow (eds.), *Meaning, use, and interpretation of language*, 302–323. Berlin: de Gruyter.
- Lohitare, L., D. Lohammarimoi, D. T. Peter & P. Joseph. 2012. *Didinga grammar book*. Juba: SIL publications.
- Lomashvili, L. & H. Harley. 2011. Phases and templates in Georgian agreement. *Studia Linguistica* 65. 233–267.
- Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania working papers in linguistics* 4(2). 201–225.
- Marantz, A. 2001. *Words*. Ms. Cambridge, MA: MIT.
- McPherson, L. & J. Heath. 2016. Phrasal grammatical tone in the Dogon languages: The role of constraint interaction. *NLLT* 34(2). 593–639.
- Müller, G. 2009. Ergativity, accusativity, and the order of Merge and Agree. In K. K. Grohmann (ed.), *Explorations of phase theory. Features and arguments*, 269–308. Berlin: de Gruyter.
- Müller, G. 2020. *Inflectional morphology in Harmonic Serialism (Advances in Optimality Theory)*. Sheffield: Equinox.
- Nevins, A. 2011. Multiple Agree with clitics: Person complementarity vs. omnivorous number. *NLLT* 29(4). 939–971.
- Noyer, R. R. 1992. *Features, positions and affixes in autonomous morphological structure*. Cambridge, MA: MIT dissertation.
- Reid, T. 2019. *The phonology and morphology of the Nuer verb*. Guildford: University of Surrey dissertation.



- Ritter, E. 1991. Two functional categories in noun phrases: Evidence from Modern Hebrew. *Syntax and semantics* 25. 37–62.
- Rolle, N. 2019. In support of an OT-DM model. *NLLT* 1–59.
- Sande, H. & P. Jenks. 2018. Cophonologies by phase. In *Proceedings of the 48th Annual Meeting of the North East Linguistic Society (NELS)*, vol. 3, 39–53.
- Sauerland, U. 2003. A new semantics for number. In R. B. Young & Y. Zhou (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) XIII*, 258–275. Ithaca: Cornell University, CLC Publications.
- Sauerland, U., J. Anderssen & K. Yatsushiro. 2005. The plural is semantically unmarked. In S. Kepser & M. Reis (eds.), *Linguistic evidence. Empirical, theoretical and computational perspectives*, 409–430. Berlin & New York: de Gruyter.
- Shlonsky, U. 1989. The hierarchical organization of subject verb agreement. Manuscript, University of Haifa.
- Siddiqi, D. 2006. *Minimize exponence: Economy effects on a model of the morphosyntactic component of the grammar*. Austin: University of Arizona dissertation.
- de Swart, H. & D. Farkas. 2010. The semantics and pragmatics of plurals. *Semantics and pragmatics* 3. 1–54.
- Thornton, A. 2019. Agreeing in number: Verbal plural suppletion and reduplication. *The Linguistic Review* 36. 531–552.
- Toweett, T. 1979. *A study of Kalenjin linguistics*. Nairobi: Kenya Literature Bureau.
- Trommer, J. 2001. *Distributed Optimality*. Potsdam: Universität Potsdam dissertation.
- Trommer, J. 2014. Ø-agreement in Turkana. In I. Bornkessel-Schlesewsky, A. L. Malchukov & M. Richards (eds.), *Scales and hierarchies: A cross-disciplinary perspective*, Berlin, Munich & Boston: de Gruyter.
- Tucker, A. N. & J. T. O. Mpaayei. 1955. *A Maasai grammar, with vocabulary*. London: Longmans, Green, and Co.
- Veselinova, L. 2006. *Suppletion in verb paradigms: Bits and pieces of a puzzle*. Amsterdam: John Benjamins.
- Wiltschko, M. 2008. The syntax of non-inflectional plural marking. *NLLT* 26.3. 639–694.
- Wood, E. 2007. *The semantic typology of pluractionality*. Berkeley: UC Berkeley dissertation.
- Zwarts, J. 2004. *The phonology of Endo: A Southern Nilotic language* (Studies in African Linguistics 59). München: LINCOM.