Grammatical gender and linguistic complexity

Volume II: World-wide comparative studies

Edited by

Francesca Di Garbo

Bruno Olsson

Bernhard Wälchli

Studies in Diversity Linguistics ??

language science press

Studies in Diversity Linguistics

Editor: Martin Haspelmath

In this series:

1. Handschuh, Corinna. A typology of marked-S languages.
2. Rießler, Michael. Adjective attribution.
3. Klamer, Marian (ed.). The Alor-Pantar languages: History and typology.
4. Berghäll, Liisa. A grammar of Mauwake (Papua New Guinea).
5. Wilbur, Joshua. A grammar of Pite Saami.
6. Dahl, Östen. Grammaticalization in the North: Noun phrase morphosyntax in Scandinavian vernaculars.
7. Schackow, Diana. A grammar of Yakkha.
8. Liljegren, Henrik. A grammar of Palula.
9. Shimelman, Aviva. A grammar of Yauyos Quechua.
10. Rudin, Catherine & Bryan James Gordon (eds.). Advances in the study of Siouan languages and linguistics.
11. Kluge, Angela. A grammar of Papuan Malay.
12. Kieviet, Paulus. A grammar of Rapa Nui.
13. Michaud, Alexis. Tone in Yongning Na: Lexical tones and morphotonology.
14. Enfield, N. J. (ed.). Dependencies in language: On the causal ontology of linguistic systems.
15. Gutman, Ariel. Attributive constructions in North-Eastern Neo-Aramaic.
16. Bisang, Walter & Andrej Malchukov (eds.). Unity and diversity in grammaticalization scenarios.
17. Stenzel, Kristine & Bruna Franchetto (eds.). On this and other worlds: Voices from Amazonia.
18. Paggio, Patrizia and Albert Gatt (eds.). The languages of Malta.
19. Seržant, Ilja A. & Alena Witzlack-Makarevich (eds.). Diachrony of differential argument marking.
20. Hölzl, Andreas. A typology of questions in Northeast Asia and beyond: An ecological perspective.
21. Riesberg, Sonja, Asako Shiohara & Atsuko Utsumi (eds.). Perspectives on information structure in Austronesian languages.

22. Döhler, Christian. A grammar of Komnzo.
23. Yakpo, Kofi. A Grammar of Pichi.
24. Guérin Valérie (ed.). Bridging constructions.
25. Aguilar-Guevara, Ana, Julia Pozas Loyo & Violeta Vázquez-Rojas Maldonado *eds.). Definiteness across languages.

ISSN: 2363-5568

Grammatical gender and linguistic complexity

Volume II: World-wide comparative studies

Edited by

Francesca Di Garbo

Bruno Olsson

Bernhard Wälchli

language science press

Di Garbo, Francesca, Bruno Olsson & Bernhard Wälchli (ed.). 2019. *Grammatical gender and linguistic complexity*: *Volume II: World-wide comparative studies* (Studies in Diversity Linguistics ??). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/237

Indexed in EBSCO
ISBN: 978-3-96110-180-1 (Digital)
978-3-96110-181-8 (Hardcover)

ISSN: 2363-5568
no DOI
Source code available from www.github.com/langsci/237
Collaborative reading: paperhive.org/documents/remote?type=langsci&id=237

Cover and concept of design: Ulrike Harbort
Fonts: Linux Libertine, Libertinus Math, Arimo, DejaVu Sans Mono, SIL Doulos, Sarabun Light, SIL Annapurna, ITF Hind Madurai
Typesetting software: XeLaTeX

Language Science Press
Unter den Linden 6
10099 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

Contents

Chapter 1

The evolving complexity of gender agreement systems

Francesca Di Garbo
Stockholm University

Matti Miestamo
University of Helsinki

This paper proposes to integrate the diachronic dimension to the typological study of gender complexity, and focuses on the morphosyntactic encoding of gender distinctions via agreement patterns. After investigating the processes of language change that foster the reduction, loss, expansion and emergence of gender agreement in a sample of fifteen sets of closely related languages (N= 36 languages), we discuss how gender agreement systems in decline and on the rise pattern in terms of complexity. We show that declining and emerging gender agreement systems may exhibit increase or decrease in complexity and discuss how this relates to the fact that they represent transitional stages between absence of gender and full-fledged gender systems. In our analysis, we make use of typological implicational hierarchies in the domain of agreement as a tool to account for diachronic variation and for the patterns of simplification/complexification in the agreement systems of the sampled languages.

Keywords: agreement hierarchy, agreement redistribution, gender emergence, gender expansion, gender loss, gender reduction, morphophonological erosion, complexification, simplification

1 Introduction and key notions

Within the last decade, pioneering research on the complexity of grammatical gender has contributed to identify a number of dimensions along which gender systems may vary in complexity (see Audring 2014; 2017; Di Garbo 2016 for

Francesca Di Garbo & Matti Miestamo. 2019. The evolving complexity of gender agreement systems. In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity: Volume II: World-wide comparative studies*, 1–43. Berlin: Language Science Press. DOI:??

gender-specific complexity measures[1]), and to apply these dimensions of complexity variation to research on the typology of gender systems within specific language families and areas of the world (Di Garbo 2016). The approach followed in these studies has been predominantly synchronic. In this paper, we argue that integrating the diachronic dimension to the typological study of gender complexity is essential to understand how gender systems vary in complexity (i.e., along which dimensions of the proposed metrics) and how this variation is distributed crosslinguistically.

We investigate the evolution of complexity in the domain of grammatical gender by using a diachronic approach to the study of linguistic diversity in line with Greenberg (1978a). Greenberg addressed possible pathways of change between different types of structures and languages and argued that there would likely be a diachronic connection between all language types in a typology in the sense that change from any given type to any other type would be possible. This diachronic route would not always be direct, but rather mediated by other types, and the relative stability of the different types would differ, with some types qualifying as *stable, persistent*, and others as *unstable, transitional*. In this paper, we describe the patterns of language change whereby complexification and simplification in gender systems take place, explore possible functional explanations to the unfolding of these changes, and show how these explanations are ultimately grounded in well-known implicational tendencies in the typology of gender systems. In addition, by operationalizing gender complexity as a dynamic, evolving variable, we explore the relationship between the complexity and stability of gender systems.[2] The questions we attempt to answer are:

- Which complexities are most stable in the domain of grammatical gender?
- Which other aspects of gender complexity are more likely to change?
- To what extent can we identify complexification or simplification in the processes of emergence and expansion of gender on the one hand, and reduction and loss of gender on the other?

[1] In addition, see Passer (2016) for a discussion of gender complexity in comparison with other nominal classification strategies; and Leufkens (2015) for a discussion of grammatical gender in the context of a general model of complexity and transparency in grammar.

[2] On the stability of gender systems see the pioneering large-scale typological investigation by Nichols (1992) as well as the more recent overview by Nichols (2003). For a study of the diachrony and stability of grammatical gender in the Indo-European family, see Matasović (2004).

Following Miestamo (2006; 2008) we define complexity in absolute, theory-oriented, objective terms, paying attention to the number of elements in a system and connections between these. In an information-theoretic perspective, complexity can ultimately be reduced to description length: of two entities, for instance two grammatical systems, the less complex one is the one whose shortest possible description is shorter. In other words, the simpler entity can be compressed into a smaller space without losing information. This approach also aligns with complexity theories outside linguistics and thereby allows linguistic complexity to be viewed in a cross-disciplinary perspective as well. The notions of cost and difficulty of processing and learning are related to complexity, and some authors, such as Kusters (2003), take a relative, user-oriented, subjective approach, equating complexity with cost and difficulty. In a user-oriented approach, those aspects of language that increase processing load and learning difficulty are defined as complex. Dahl (2004) and Miestamo (2006; 2008) discuss some obvious problems with the cost- and difficulty-based approach and point out that it is important to keep the notions of complexity and difficulty apart. However, to what extent and in what ways complexity and difficulty are correlated is a highly interesting question. We believe that keeping these notions apart is a prerequisite for adequately addressing this issue.

Miestamo (2006; 2008) proposes two principles by which grammatical complexity can be measured:

- The Principle of Fewer Distinctions, which, paying attention to grammatical meaning, defines as less complex a grammatical system in which, other things being equal, fewer semantic/pragmatic distinctions are made grammatically.
- The Principle of One-Meaning–One-Form, which, paying attention to the relationship between meaning and form, defines as less complex those systems and structures in which, other things being equal, each meaning is expressed by one form and each form corresponds to only one meaning.

Violations of these two principles increase complexity.

To take some examples, by the Principle of Fewer Distinctions, a gender system with two grammaticalized gender distinctions is less complex in this respect than a gender system with, say, five grammaticalized distinctions. By the Principle of One-Meaning–One-Form, we can identify a higher degree of complexity in a gender system system in which: (a) the formal expression of one or more genders is combined with other categories in one morpheme (fusion, multiple exponence);

(b) one or more gender distinctions are expressed with multiple/discontinuous morphemes (fission); (c) the markers of one or more gender distinctions show two or more variants (allomorphy); and/or (d) the markers of some gender distinctions are identical in some grammatical contexts (syncretism).

While the Principle of One-Meaning–One-Form can handle the relation between meaning and form relatively exhaustively (relevant subcriteria need of course to be defined and refined), the Principle of Fewer Distinctions only covers parts of complexity on the level of meaning. Things get more complicated when we look at the interaction between different functional domains (e.g., gender and number). Dahl (2004) discusses the notion of choice structure, i.e. the dependency of available choices on choices made earlier (cf. also the notion of dependency hierarchies by Aikhenvald & Dixon 1998). To take an example from the domain of grammatical gender, in many languages gender distinctions are available only in the singular, but are neutralized in the plural. This is, for instance, the case in Russian (Indo-European, Slavic). In order to account for interactions between functional domains and their effect on the complexity of individual domains, Di Garbo (2014; 2016) proposes the Principle of Independence.

- The Principle of Independence defines as less complex those systems and structures which, other things being equal, are *independent* of other systems and structures.

Under the Principle of Independence, a gender system whose formal realization is dependent on number distinctions is more complex than a gender system which is not constrained by number distinctions.

The three principles, the Principle of One-Meaning–One-Form, the Principle of Fewer Distinctions, and the Principle of Independence, are all operationalized in the gender complexity metric proposed by Di Garbo (2014; 2016), as well as in the discussion of gender complexity and canonicality by Audring (2019 [in Volume I]).[3] In this paper, we will be especially concerned with the way in which morphosyntactic and semantic properties of reducing and emerging gender systems may be accounted for as violations of one of these principles.

The paper is organized as follows. §2 presents some of the parameters along which gender systems may vary, and the sampling method followed in the study. In §3, attention is given to the factors that explain synchronic variation in the domain of gender agreement and to the extent to which these can be mapped on

[3] Audring (2019 [in Volume I]) uses a different terminology for the Principle of One-Meaning–One-Form and the Principle of Fewer Distinctions. In her own terminology, these are the Principle of Transparency and the Principle of Economy, respectively.

diachronic change, too. Reducing gender agreement systems are presented in §4 whereas §5 focuses on emerging gender agreement systems, and §6 on expanding gender agreement systems. In §7, we discuss how changes in the domain of gender agreement affect the complexity of gender systems. Concluding remarks are given in §8.

2 The evolution of gender complexity

In this paper, we explore synchronic distributions of types of gender systems among closely related languages, and, based on these synchronic distributions, we try to infer how gender systems change through time becoming more or less complex. We draw our observations from a sample of fifteen language sets. Each set consists of two to three genealogically related languages. In addition, the sample includes one isolate within the Austronesian family, Chamorro, and one mixed language, Michif. The total number of languages is 36. The map in Figure 1 illustrates the geographic distribution and genealogical affiliations of the sampled languages. A list of the sampled languages can be found in Appendix 8.

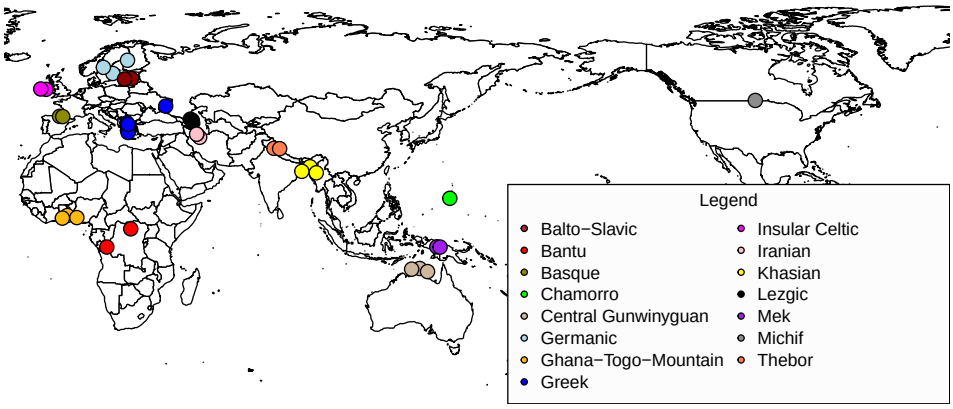

Figure 1: The language sample

The data set studied stems from a larger project on the sociohistorical correlates of the evolution of gender complexity led by Francesca Di Garbo (for details, see Di Garbo forthcoming). The diachronic processes examined in the study are somewhat biased towards instances of contact-induced change, even though language-internal developments are also discussed. While the pace and nature of these developments may thus be specific to the type of contact situation in which they unfold, we believe that the data set under study offers insights of

rather general relevance with respect to the diachrony of gender marking systems. Data were collected based on a questionnaire (Di Garbo 2015), as well as on consultation of reference grammars and language experts.

Typological research on grammatical gender systems has mostly focused on three broad domains of analysis:

- Number of genders
- Number and/or type of gender assignment rules
- Formal marking through agreement patterns.

We argue that these domains of synchronic variation can also be used to investigate how gender systems change through time. However, we suggest that any change in the number of gender values or the number and nature of gender assignment rules must ultimately hinge on variation and change in the domain of agreement patterns, that is, in the morphosyntactic encoding of gender distinctions. For instance, a gender value is lost when the corresponding gender agreement patterns fall out of use. Similarly, changes in the nature and distribution of gender assignment rules are reflected by the gender agreement patterns that the nouns affected by these changes trigger in discourse. For instance, we know that a former masculine noun is re-analyzed as neuter if patterns of neuter agreement are selected when the noun is used. Thus, we argue that studying synchronic and diachronic variation in patterns of gender agreement enables us to make generalizations about variation and ongoing change in the number of genders and/or the nature of the gender assignment rules that languages have. This suggestion aligns with recent observations in the literature on gender complexity where complexity in the domain of gender agreement has been shown to interact with complexity at the level of gender values and assignment rules (Audring 2017; Di Garbo 2016).[4]

We explore simplification and complexification of gender systems by focusing on reducing, emerging and expanding patterns of gender agreement. The sample languages are thus selected so as to represent instances of (1) reduction, (2) loss, (3) emergence, and (4) expansion of gender agreement. These are then compared with instances of retention or lack of gender agreement as attested in closely related languages. Naturally, loss, reduction and expansion presuppose the preexistence of a gender system within the relevant language sets, whereas emergence of gender presupposes absence of gender within the relevant language sets.

[4]For instance, Di Garbo (2016) shows that manipulable gender assignment tends to presuppose rather pervasive gender agreement systems in the languages of her sample.

The data in Table 1 and the map in Figure 2 illustrate how the patterns of change in focus are distributed within the languages of the sample.[5]

Table 1: Patterns of change attested in the languages of the sample

Family by macroarea	Language	Pattern of change
Eurasia		
Khasian	Khasi	Expansion
	Lyngngam	Retention
	Pnar	Expansion
Basque	Standard Basque	Lack
	Lekeitio Basque	Emergence
Balto-Slavic	Latvian	Retention
	Tamian Latvian	Loss
Greek	Modern Greek	Retention
	Pontic Greek	Reduction
	Rumeic Greek	Reduction
	Cappadocian Greek	Loss
Insular Celtic	Irish	Reduction
	Irish (Ros Much)	Retention
North Germanic	Elfdalian	Retention
	Karleby Swedish	Reduction
	Standard Swedish	Reduction
Northwestern Iranian	Eshtehardi	Expansion
	Kafteji	Expansion
	Kelasi	Loss
Lezgic	Archi	Retention
	Aghul	Loss
	Udi	Loss
Thebor	Shumcho	Emergence
	Jangshung	Emergence
Papunesia		
Chamorro	Chamorro	Emergence
Mek	Nalca	Emergence
	Eipo	Emergence

[5]For language classification we follow the Glottolog (Hammarström et al. 2018).

Family by macroarea	Language	Pattern of change
Africa		
Bantu	Kinshasa Lingala	Reduction
	Makanza Lingala	Expansion
Ghana-Togo-Mountain	Selee	Retention
	Igo	Reduction (near loss)
	Ikposo	Loss
Australia		
Gunwinggu	Kunwinjku	Retention
	Kundjeyhmi	Reduction
	Kune	Loss
North America		
Mixed Language	Michif	Expansion

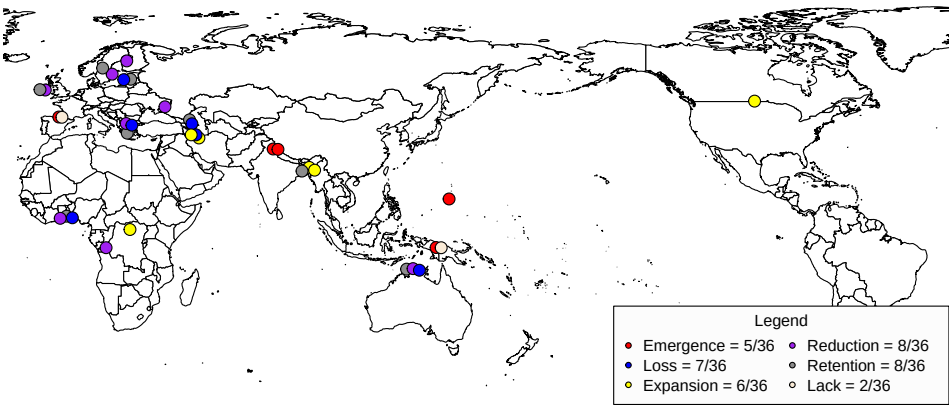

Figure 2: Distribution of patterns of change

It can be hypothesized that gender agreement systems in decline represent instances of reducing complexity, while gender agreement systems on the rise or under expansion represent instances of increasing complexity. A further possible hypothesis is that gender agreement systems on the rise or in decline are less complex than the more pervasive systems that they are moving towards or away from. We will come back to these hypotheses in §7 and evaluate them against our data.

3 The evolution of gender complexity in the domain of agreement

Starting with the pioneering work by Corbett (1979; 1991), a great deal of research has focused on unraveling constraints on the distribution of gender distinctions on different types of agreement targets. This research has shown that certain agreement targets (e.g., personal pronouns) are more likely than others (e.g., attributive modifiers) to index semantic rather than grammatical properties of nouns. In the terminology proposed by Corbett (1979; 1991), this is known as an opposition between *semantic* and *syntactic* agreement patterns. Preferences towards semantic or syntactic agreement per type of agreement target are captured in the form of an implicational hierarchy, which is known as the Agreement Hierarchy. The Agreement Hierarchy – illustrated in (1) – was first proposed by Corbett (1979) and is further discussed in Corbett (1991; 2000; 2006). It expresses the likelihood of semantic agreement to occur with different types of agreement targets as well as the degree of syntactic cohesion between agreement targets and their controllers.

(1) The Agreement Hierarchy (adapted from Corbett 2010)

- Semantic Agreement
  attributive > predicate > relative pronoun > personal pronoun
- Syntactic Cohesion
  attributive < predicate < relative pronoun < personal pronoun

The directions of the arrows – ">" or "<" – stand for different directionalities in the two main chains of implications entailed by the hierarchy. The first row indicates that semantic agreement on any of the targets to the left implies the presence of semantic agreement on the targets to the right, with attributive modifiers being the least likely candidate for semantic agreement. The second row indicates that syntactic cohesion between nouns and any of the targets to the right of the hierarchy implies at least the same level of syntactic cohesion with any of the targets to the left, with personal pronouns being the agreement targets with the loosest syntactic integration to nouns. These hierarchical effects are connected with the fact that pronouns tend to be linearly more distant from their antecedents (low syntactic cohesion) as compared, for instance, with definite articles (high syntactic cohesion), which tend to occur linearly closer to the controller nouns.[6] Pronouns are therefore more prone to index semantic proper-

[6] Different types of agreement targets may occur within the noun phrase (articles, quantifiers, numerals etc.) and further hierarchical effects between such targets cannot be excluded. This, however, falls outside the scope of the present investigation.

ties of the discourse referent rather than lexico-grammatical properties of nouns, such as grammatical gender. Mismatches between the agreement patterns associated with different types of targets are especially likely to occur when the controller nouns are *hybrid nouns*. In the case of gender, these are nouns whose inherent gender assignment is in conflict with their semantics. A classic example is the German noun for 'girl', *Mädchen*, which is grammatically neuter, but denotes a human entity. Let us consider the types of gender agreement mismatches attested in German with the noun *Mädchen*.

(2) German (Indo-European, Germanic; Corbett 1991: 228)
Schau dir dieses Mädchen an, wie gut sie/es Tennis spielt
look you this.N girl at, how good she/it tennis plays
'Look at this girl, see how well she plays tennis.'

The example shows that while gender agreement within the noun phrase (i.e., on the demonstrative) can only conform to the lexical gender of the noun (*dieses*, N), speakers can choose between feminine and neuter agreement for personal pronouns. Feminine agreement indexes the fact that the discourse referent is female (as in *sie*, F); neuter agreement indexes the fact that the noun for 'girl' is grammatically neuter (as in *es*, N).[7] Conflicts between "semantic" and "syntactic" agreement can also be understood in terms of mismatches between *referential* and *lexical* gender, as these terms are used by Dahl (2000) (see also the study of the evolution of gender marking in medieval English by Siemund & Dolberg 2011).

replaced 'with the personal pronouns, speakers can choose between feminine and neuter agreement.'

There are at least two ways in which the Agreement Hierarchy can be used to describe synchronic variation in gender complexity, one pertaining to the types and number of attested agreement domains, and one pertaining to the type and number of preferred agreement patterns per domain. Concerning type and number of attested agreement domains, a language that exhibits gender agreement in all the agreement domains represented along the hierarchy is, in this respect, more complex than a language that, other things being equal, has agreement in fewer domains. This is, for instance, the way in which the amount of gender agreement or gender indexation is treated in the metric proposed by Di Garbo (2016).[8] Concerning type and number of preferred agreement patterns, a lan-

[7] Corbett (1991: 228) further mentions that the older the age of the young woman that is being talked about, the more likely it is for speakers to use feminine agreement.

[8] For some observations on possible implicational tendencies constraining which agreement domains are more likely to be targets of gender marking in a sample of 20 languages from New Guinea see Svärd (2019 [in Volume I]).

guage in which gender agreement is only syntactic with all agreement targets is, in this respect, less complex than a language that, other things being equal, exhibits variation between syntactic and semantic agreement at any point along the hierarchy. For a broader discussion about the use of typological implicational hierarchies as cross-linguistic measures of complexity, see Miestamo (2009).

In this paper, we explore the extent to which not only synchronic, but also diachronic variation in the domain of gender agreement can be mapped onto the Agreement Hierarchy (for an overview of the role of the Agreement Hierarchy in the diachrony of nominal classification see also Seifart 2010). With respect to types and number of agreement domains, we find that, in the languages of our sample, both the rise and the decline of gender agreement tend to start off from the agreement domains at the two opposite ends of the Agreement Hierarchy, i.e., either from attributive modifiers or from personal pronouns and/or other type of anaphoric constructions, such as light nouns with anaphoric functions (for the latter, see also Wälchli 2019 [this volume]). With respect to types and number of preferred agreement patterns per domain we find that, in the languages of the sample, at least the decline and loss of gender agreement tend to be directional, and that the attested lines of directionality are reminiscent of the two opposite pulling forces described by the Agreement Hierarchy: syntactic cohesion between controllers and targets, and spread of semantic agreement. However, we make no claims about the universality of these tendencies, and we do not exclude that, in languages other than those sampled for this study, diachronic change in the morphosyntax of gender agreement occurs on other types of agreement targets first. Finally, while we argue that the hierarchy is a useful tool to *describe* tendencies in how gender marking systems change, we make no claims about it having a *predictive/explanatory* value concerning the spreading of such changes. On the contrary, we argue that explanations should be sought in the realm of those functional pressures that are reflected in the hierarchy.

In §4, we focus on reducing gender agreement systems; emerging gender agreement systems are discussed in §5 whereas the expansion of gender agreement patterns is treated in §6.

4 Reducing gender agreement systems

4.1 Attested processes of change

In our data, the reduction and, in some cases, the loss of gender agreement result from two distinct diachronic processes: (1) *morphophonological erosion* and (2) *redistribution* of agreement patterns.

By morphophonological erosion we refer to the wholesale patterns of change that lead to the loss of inflection. Sound changes (e.g., changes in stress patterns resulting in the loss of word-initial or word-final segments) can cause loss of segmental morphology, which ultimately determines the neutralization of previously overtly coded grammatical distinctions and the overall restructuring of inflectional paradigms. This process is also known in the literature under the label *deflection*. Within the domain of nominal morphology, morphophonological erosion often affects gender marking along with the marking of other nominal inflectional features, such as number and case, which are frequently cumulatively encoded with gender. It has been suggested (see Priestly 1983 for Indo-European; Audring 2009 for Germanic languages) that, when morphophonological erosion affects the encoding of gender distinctions, the word classes that are likely to lose gender marking first are the nouns themselves (in case of overt gender systems), followed by the agreement targets that are more adjacent to nouns, i.e., adnominal modifiers, such as definiteness markers, demonstratives, adjectives and numerals, with definiteness markers generally being yet more stable than, say, numerals or adjectives. Personal pronouns (both dependent and independent) are more likely to retain the encoding of gender distinctions as a means to signal semantic properties of the discourse referents. In other words, under morphophonological erosion, gender agreement is more likely to be retained on those agreement targets where it is most functional to reference tracking and reference identification, i.e. demonstrative and/or personal pronouns. These may then tend to inflect based on semantically transparent principles of gender assignment (animacy and/or biological gender). In English, for instance, the encoding of gender distinctions underwent massive erosion as part of a general weakening of inflectional morphology. As a result of this deflection process, gender marking was lost on all of the agreement targets (as well as on nouns) except for the personal pronouns, which nowadays signal the biological gender of discourse referents, and for the relative pronouns which make a distinction of the human/non-human type (Curzan 2003).

By *redistribution of agreement*, we refer to the process whereby one of the several agreement patterns available in a language (for instance, the neuter) starts being used with nouns that would normally trigger agreement in other genders (for instances, with nouns that are semantically inanimate, but grammatically masculine or feminine). If the redistribution of one agreement pattern comes to affect all agreement domains, and to effectively replace all the other competing

agreement patterns independently of semantic or morphological properties of the controller nouns, then gender distinctions become neutralized. In many of the cases attested in our sample, the redistribution of agreement patterns appears to be at least initially semantically motivated: semantic oppositions generally pertaining to the domain of animacy start affecting the criteria according to which certain nouns trigger gender agreement on at least some targets. In general, the higher the number of nouns involved in the restructuring of the assignment criteria, the higher the chance that the overall gender assignment rules of a language may change. Similarly, the higher the number of agreement targets that align with the new assignment criteria, the more reasons to speak of an increase or decrease in the number of gender distinctions. For instance, when the semantic agreement patterns that are being redistributed are based on animacy, their generalization to all agreement targets may eventually lead to a bipartite, animate vs. inanimate, type of gender system, where gender assignment is semantically predictable. This is for instance the case of the Bantu language Kinshasa Lingala, in which all productive agreement targets index the animacy of the noun, whereas the nouns themselves retain prefixal remnants of the old, no longer productive system of gender distinctions (Maho 1999: 130–132; Meeuwis 2013: 28–29). In other cases, the most frequent (default) pattern of gender agreement is the one that takes over. This is for instance the case of Tamian Latvian (Indo-European, Balto-Slavic), where the masculine agreement pattern has replaced nearly all instances of feminine agreement leading to loss of grammatical gender. The redistribution of agreement patterns is ultimately a process of analogical levelling: the gender agreement system of a language is restructured on the basis of the more semantically motivated and/or more frequent agreement pattern, which gradually spreads at the expenses of others.

Table 2 illustrates the distribution of patterns of reduction and loss of gender agreement within the languages of the sample, and specifies whether these are due to morphophonological erosion, redistribution of agreement, a combination of both, or whether the exact pattern of change cannot be inferred based on the data at our disposal. For each of the relevant languages, the table also specifies if directionality applies, and if the distribution of a given pattern of change is at any rate semantically motivated. Given the limited size of our sample, the analysis proposed here is merely qualitative and we draw no generalization based on the relative frequencies of the observed patterns of change. Examples for each of the possible scenarios are discussed in §4.2, 4.3 and 4.4.

Table 2: Morphophonological erosion and redistribution of agreement in the languages of the sample where gender agreement reduction and loss are attested

	Languages	Directionality	Semantics
Morphophonological erosion	Standard Swedish	YES	NO
	Kelasi	Not clear	NO
Redistribution	Cappadocian Greek	YES	YES
	Pontic Greek	YES	YES
	Rumeic Greek	YES	YES
	Irish	YES	YES
	Kune	Not clear	No data
Both	Igo	YES	Not clear
	Karleby Swedish	Not clear	Partially
	Kinshasa Lingala	YES	YES
	Tamian Latvian	Partially	Partially
Not clear	Aghul	–	–
	Kundjeyhmi	–	–
	Lezgian	–	–
	Udi	–	–

Table 3: Personal Pronouns in Standard Swedish

	M	F	PL
Nominative	*han* 'he'	*hon* 'she'	*de* 'they'
Genitive	*hans* 'his'	*hennes* 'her'	*deras* 'their'
Accusative	*honom* 'him'	*henne* 'her'	*dem* 'them'

4.2 Reduction and loss by morphophonological erosion

In Standard Swedish, the opposition between masculine and feminine gender is retained in the inflectional paradigm of the independent third person pronouns (see Table 3), but has been lost elsewhere.[9]

The Masculine and Feminine singular forms of the third person pronouns are used to signal the biological gender of human and other animate referents.[10]

[9] In written language, a masculine suffix *-e* may still sometimes be used on adjectives to mark masculine agreement.

[10] During the last decade, a biological gender-neutral form, *hen* has been introduced. Its frequency of use has rapidly increased, both in written and spoken Swedish discourse.

With non-animate entities, the demonstrative pronouns *den*, Common Gender, and *det*, Neuter Gender, are used instead, and the choice between the two is based on the lexical gender of nouns. In sum, in the pronominal domain, Standard Swedish has a four-way gender distinction: Masculine, Feminine, Common, Neuter, with a split between animate and inanimate referents governing the distribution of these gender values. Within the domain of adnominal modification, Swedish distinguishes between a Common and a Neuter Gender only: *en person* 'a person' (Common Gender), and *ett hus* 'a house' (Neuter Gender). Historically, the Common Gender is the result of a merger between the Feminine and Masculine genders. Many nonstandard varieties of Swedish, as well as many other Scandinavian varieties, retain a tripartite gender system. Tripartite gender systems were found all over Scandinavia before the standard varieties with a bipartite gender system, such as Danish and Swedish, started spreading.[11] One of the Swedish dialects which still retains a fully productive tripartite gender system is Elfdalian, spoken in the Swedish region of Northern Dalarna by approximately two thousand people.[12] In Elfdalian, the opposition between Masculine, Feminine and Neuter gender runs productively through the whole agreement system. A tripartite gender system of the type retained by Elfdalian is also attested in Old Swedish texts.[13] The Masculine-Feminine merger in the domain of adnominal modification appears to be due to a combination of various morphophonological processes, such as the erosion and loss of the masculine *-er* ending in the inflectional paradigm of strong adjectives, the loss of the masculine suffix *-r* before the definite suffix in the nominative form of the noun, and the loss of final consonant length in the inflectional paradigm of the definite suffixes (Duke 2010: 652–654). Finally, pervasive reduction in gender agreement domains is attested in Karleby Swedish, the variety of Swedish spoken in the town of Karleby, located in the Finnish region of Ostrobothnia.[14] Gender agreement reduction in Karleby Swedish is best described as an instance of both morphophonological erosion and agreement redistribution. It is therefore discussed in §4.4.

[11] Before the spread of the standard languages, bipartite gender systems were only attested in Denmark, southern Sweden, the Mälaren valley in Sweden, and pockets of Norway where varieties heavily influenced by Danish were spoken (Östen Dahl, personal communication).

[12] Data from Åkerberg (2012), as well as from Östen Dahl (personal communication).

[13] The use of the Masculine and Feminine pronouns with inanimate antecedents continued in the written language until the nineteenth century, even though this distinction was lost in all other domains of nominal inflection and no longer maintained in spoken use (Östen Dahl, personal communication).

[14] It is worth mentioning that, contrary to Karleby Swedish, some other Ostrobothnian varieties of Swedish display quite conservative gender systems (for more details see Huldén 1972: 40–50.) However, it is perhaps unsurprising, that the near loss of gender distinctions is attested in the northernmost corner of the Swedish speaking area of Finland.

Loss of gender in Kelasi, a Northwestern Iranian language of the Tatic subbranch, is also the result of a process of morphophonological erosion. Stilo (to appear) proposes a historical-comparative analysis of gender loss in Kelasi whereby the decline of gender marking is explained as originating from the domain of noun inflection. In Kafteji, a closely related language spoken at a distance of twelve kilometers from Kelasi, gender distinctions are still retained. However, in Kafteji, overt marking of gender on nouns is dropped when nouns are used in a generic sense or as citation forms, and gender is never marked on agreement targets when these occur in isolation. Based on this comparative evidence, Stilo (to appear: 27) hypothesizes that, at some point in the history of Kelasi, gender marking became increasingly optional and "went through gradual stages of erosion by becoming more and more rarely used in speech", to be finally dropped in all domains of encoding. Even though the individual stages of this process of erosion are not known, nouns – "the crucial locus of gender in the grammar" of Kelasi (Stilo to appear: 27) – are viewed as the word class from which the decline of gender marking originated. This is why we classify Kelasi as an instance of gender loss by morphophonological erosion.

The reduction and loss of gender inflections as a result of a more general erosion of nominal morphology are widely attested across different genera of the Indo-European language family. See Audring (2009: chapter 9) for an overview of patterns of gender reduction and loss across Germanic languages; Priestly (1983) for a broader overview of the Indo-European language family, and, in particular, of pronominal relics of the neuter gender in Romance (e.g., Italian, French) and Baltic (e.g., Lithuanian) languages.

4.3 Reduction and loss by redistribution of agreement

Gender reduction and loss as a result of the redistribution of agreement patterns are widely attested in our sample. In this section, we discuss a selection of the attested cases.

The Asia Minor Greek dialects are a group of Greek varieties that are or, prior to the 1923 population exchange between Greece and Turkey, used to be spoken in Turkey. Karatsareas (2014) identifies five main dialects within the Asia Minor Greek cluster: Cappadocian, Pharasiot, Pontic, Silliot, and Rumeic. While the first four varieties were spoken in different areas of modern Turkey, Rumeic is the variety spoken by the Greek inhabitants of Mariupol, Ukraine, and can be considered as the historical descendant of the Pontic spoken by Greek settlers in Crimea.

Due to their long-lasting history of isolation from mainland varieties of Greek, and, partially, to a history of prolonged contact and bilingualism with Turkish, the Asia Minor Greek dialects exhibit a wealth of grammatical innovations among which a significant reorganization of the gender agreement and gender assignment patterns. This is attested in all Asia Minor Greek varieties but Silliot, which rather retains a conservative system similar to the one attested in Standard Greek and in other Modern Greek varieties outside the Asia Minor area (Karatsareas 2014: 83). Examples (3), (4), (5), and (6) illustrate the innovations attested in the domain of gender agreement and gender assignment in four out of the five groups of Asia Minor Greek dialects. We present data from the dialects that display renewed gender systems and compare them with equivalent structures in Standard Greek, where these innovations are not attested.[15]

In Pontic, example (3), the inanimate feminine noun for 'door' triggers neuter agreement with agreement targets non-immediately adjacent to nouns. In the corresponding Standard Greek sentence, agreement is feminine with all targets.

(3) a. Argyroúpolis Pontic (Indo-European, Greek; Karatsareas 2014: 79)
i *pórta* *(…) móno ímoson* *óran* *estéknen*
DEF.F.SG door.F.SG (…) only half.N.SG hour.F.SG stay.PST.3SG
anixtón
open.N.SG
'The door would stay open for only half an hour'.

b. Standard Greek (Indo-European, Greek; Karatsareas 2014: 80)
i *pórta* *móno misí* *óra* *émene* *anixtí*
DEF.F.SG door.F.SG only half.F.SG hour.F stay.PST.3SG open.F.SG
'The door stayed open for only half an hour.'

In Pontic, the criteria of gender assignment are reorganized based on the animacy of the noun: semantically inanimate, but grammatically masculine and feminine nouns are to a large extent treated as neuter. This semantic reorganization is reflected at the level of agreement: semantic (neuter) agreement with inanimate masculine and feminine nouns is attested on all agreement targets but prenominal definite articles, which instead agree with the grammatical gender of the nouns (i.e. they take masculine or feminine inflection).

In Rumeic, example (4), the pattern of semantic agreement observed in Pontic is generalized to all targets: the inanimate noun for 'winter' (which is masculine in Standard Greek) triggers neuter agreement with all agreement targets.

[15] Notice that the Standard Greek examples reported by Karatsareas (2014) can be either full or partial translations of the corresponding example in one of the Asian Minor Greek dialects.

(4) a. Rumeic (Indo-European, Greek; Karatsareas 2014: 79)
tu ko mas to ʃumós en xlísku
DEF.N.SG POSS.N.SG 1PL.GEN DEF.N.SG winter.N.SG be.PRS.3SG tepid.N.SG
'Our winter is tepid'.

b. Standard Greek (Indo-European, Greek; Karatsareas 2014: 80)
o ðikós mas o çimónas
DEF.M.SG POSS.M.SG 1PL.GEN DEF.M.SG winter.M.SG
'our winter'

In Rumeic, the gender system has been restructured based on semantic grounds: male entities are assigned to the Masculine Gender, female entities to the Feminine and inanimate entities to the Neuter.

A different path is taken by Pharasiot and Cappadocian, where the redistribution of the neuter gender agreement pattern leads to a more pervasive erosion of the gender system. In Pharasiot, as illustrated in example (5), the animate noun for 'woman' (feminine in Standard Greek) triggers neuter agreement with all targets but the definite article adjacent to the noun.

(5) a. Pharasiot (Indo-European, Greek; Karatsareas 2014: 79)
férinke adzíno i néka xortáre
bring.PST.3.SG DEM.DIST.N.SG DEF.F.SG woman.F.SG herb.PL
'that woman used to bring herbs.'

b. Standard Greek (Indo-European, Greek; Karatsareas 2014: 80)
ecíni i jinéka
DEM.DIST.F.SG DEF.F.SG woman.F.SG
'that woman'

In Pharasiot, the neuter agreement has been generalized to all nominal types (animate and inanimate) and the semantic opposition between animate and inanimate entities has been neutralized. Only the agreement targets that are most adjacent to nouns retain agreement with the original grammatical gender of the noun (in this case with the Feminine).

Finally, in Cappadocian, example (6), the neuter agreement pattern is generalized to all nouns, irrespective of animacy and type of target (the noun for 'wall' is masculine in Standard Greek).

(6) a. Axó Cappadocian (Indo-European, Greek; Karatsareas 2014: 79)
t *spitçú* *ta* *ndix(u)s xtizména*
DEF.SG.GEN house.SG.GEN DEF.PL wall.PL built.PL
'The walls of the house (are) built.'

b. Standard Greek (Indo-European, Greek; Karatsareas 2014: 80)
i *tíçi* *ine* *xtixméni*
DEF.M.PL wall.M.PL be.PRS.3PL built.M.PL
'the walls are built'.

In Cappadocian, pervasive redistribution of the neuter agreement pattern has led to complete gender loss, whereby agreement patterns only index number distinctions, in this case that the noun is plural.[16]

Using internal reconstruction, historical data, and data from contemporary varieties of Pontic spoken in Greece, Karatsareas (2014) shows that two main orders of facts account for the rise and spread of semantic agreement in Pontic. On the one hand, the triggers of semantic agreement are nouns at the bottom of the Individuation Hierarchy (Sasse 1993), that is, inanimate mass and abstract nouns that are grammatically assigned to the masculine or feminine genders. These are typical instances of hybrid nouns, i.e., nouns whose denotational semantics is in conflict with their grammatical gender assignment (these nouns denote inanimate entities, but are grammatically masculine or feminine). On the other hand, according to Karatsareas' reconstruction, the spreading of semantic agreement starts from the personal (and demonstrative) pronouns. In Pontic, the sole agreement targets that are left untouched by these redistribution patterns are those that are most adjacent to nouns, i.e., prenominal definite articles. Rumeic is the only Asia Minor Greek dialect where semantic agreement has become generalized to all nouns and targets leading to a gender system which is still tripartite (Masculine, Feminine, Neuter), but in which assignment rules and agreement patterns are entirely semantic. Conversely, in Pharasiot and Cappadocian, the generalization of the neuter agreement pattern to human nouns has paved the way for a more pervasive erosion of gender marking.[17] This process of erosion has turned into complete loss in (varieties of) Cappadocian only. The loss of gender in Cappadocian Greek is seen by Karatsareas (2014: 99) as reasonably connected

[16] Feminine and masculine agreement survive in the singular form of definite articles preceding nouns only in the Delmesó, Potámia, and Sílata varieties of Cappadocian (Karatsareas 2014: 97).

[17] A similar development is attested in some more recent varieties of Pontic, where at least human nouns denoting female referents systematically trigger neuter agreement (Karatsareas 2014: 96–97).

with the fact that, among all Asia Minor Greek varieties, this is the one with the longest and tightest history of contact and bilingualism with Turkish. A summary of the patterns of agreement redistribution attested in Pontic, Rumeic, and Cappadocian Greek is given in Figure 3.

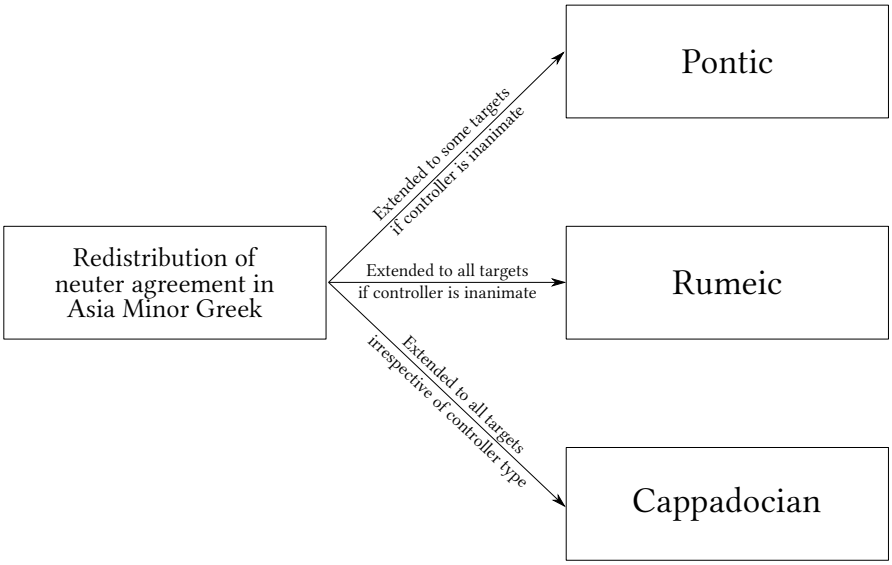

Figure 3: Neuter Agreement in the Asia Minor Greek dialects

Semantically motivated redistribution of gender agreement patterns also occurs in contemporary varieties of urban Irish as documented by Frenda (2011). In these non-standard varieties of Irish (which Frenda classifies as "non-native"), masculine agreement is increasingly used as the default agreement pattern for grammatically feminine nouns denoting inanimate entities. The redistribution is very pervasive in the domain of personal pronouns where the gender assignment system appears to be largely based on an opposition between "female referent" (marked by the Feminine Pronoun) and "everything else" (marked by the Masculine Pronoun). In the domain of adnominal modification, controller nouns that are grammatically feminine but semantically inanimate still trigger feminine agreement (this is attested in 88% of the examined cases; see Frenda 2011: 17, Figure 1).

In sum, the data from our sample suggest that patterns of agreement redistribution tend to be constrained by the syntactic cohesion between controller nouns and agreement targets. Those agreement targets that are most adjacent to nouns are the ones that are affected last by the spreading of innovations.

4.4 Combined and unclear cases

In some cases, both morphophonological erosion and agreement redistribution are attested in one and the same language, albeit not necessarily as the result of co-occurrent patterns of change. One such case is Igo, a Ghana-Togo-Mountain language of the Kwa subfamily of the Atlantic-Congo family, spoken by approximately 6.000 people (Gblem-Poidi 2007). In general, the Ghana-Togo-Mountain languages represent an ideal test case for an intragenealogical study of the diachrony of gender systems and their evolving complexity (for a historical-comparative overview, see also the contribution by Güldemann & Fiedler 2019 [in Volume I]). Some languages within the family, such as Selee (Agbetsoamedo 2014) and Siwi (Dingemanse 2009), display very productive gender systems characterized by a high number of (non-sex-based) gender distinctions, pervasive agreement and overt marking of gender on nouns. Some other languages (e.g., Animere) present heavily eroded and completely semanticized systems of gender assignment and gender agreement, whereby gender assignment and agreement are animacy-based, and traditional noun class marking on nouns is retained merely as a means of marking singular/plural distinctions. Finally, a few other languages, such as Ikposo (Soubrier 2013), have lost gender completely and retain relics of the extinct gender marking system only on nouns. Igo provides us with an example of a system in transition from animacy-based gender distinctions (of the Animere type) to complete loss of gender (of the Ikposo type). Gblem-Poidi (2007) argues that the original gender system of Igo consisted of eleven non-sex-based genders whose distribution paralleled the eleven pairings of singular and plural nominal prefixes still in use in the language. Nowadays, however, in formal registers of Igo,[18] only an animate/inanimate type of distinction is marked on the agreement targets. It can thus be assumed that this animacy-based gender system is already an eroded system, and that this process of erosion may have occurred through the spreading of semantic, animacy-based agreement. Albeit preferred in formal registers and still in use among the older generations, the animacy-based gender system of Igo is described by Gblem-Poidi (2007) as under threat, highly eroded in the speech of middle-aged speakers, and practically unused by the younger speakers. The ongoing loss of gender distinctions in Igo is the result of the erosion of segmental gender morphology. Gender agreement morphemes are omitted in actual discourse while their tonal patterns are retained in the form of floating tones that encroach upon the immediately fol-

[18] Those in use in the literacy program and in the New Testament Translation (Honorine Gblem-Poidi, personal communication).

lowing tonal segments. Interestingly, in spoken use, the former animate gender agreement markers (*ù-* and *bù-*) are resumed and reanalyzed as nominal number markers, whereby *ù-* marks the singular with both animate and inanimate nouns, and *bù-* the plural, but only with animate nouns. Example (7) shows overt plural marking with animate nouns and zero marking with inanimate.

(7) Igo (Niger-Congo, Kwa, Ghana-Togo-Mountain; Gblem-Poidi 2007: 59)

a. *bégù lɔ̄ bù ɖā wūlū*
children DEF PL PROG.SBJ dry.out
'The children are losing weight'.

b. *ātī lɔ̄ ɖàā wūlū*
trees DEF PROG.SBJ dry.out
'The trees are dying out'.

Based on the data at our disposal, it is not possible to determine whether the loss of segmental gender marking affects all agreement targets at once or is gradually spreading from one agreement domain to the other.

Another instance of pervasive reduction of gender agreement morphology which seemingly results from a combination of morphophonological erosion and agreement redistribution is Karleby Swedish. In this variety of Swedish, gender distinctions have been lost on all agreement targets except for the definite articles (immediately adjacent to nouns) and the demonstrative and personal pronouns. These retain a tripartite distinction between Masculine, Feminine and Neuter gender. The masculine and feminine forms are however used only when the controller noun denotes human beings; in all other cases only one form (the Neuter) is used both in the domain of definite and indefinite articles and with demonstrative and anaphoric pronouns (Huldén 1972; Hultman 1894). It is reasonable to think that this superimposed animacy-based distinction (whereby only nouns denoting humans trigger a masculine/feminine distinction) might have spread from the domain of anaphoric pronouns (where, for instance, it is also found in Standard Swedish) to the definite articles.

In the Tamian dialects of Latvian, loss of gender marking is also the result of a complex interplay between morphophonological erosion and agreement redistribution. According to the recent comparative study by Wälchli (2017), the loss of short vowels in final syllables caused the neutralization of the opposition between masculine and feminine gender in the accusative plural of nominal paradigms. The neutralization pattern later extended to the demonstratives. This paved the way to several processes of redistribution that led to the gradual generalization of masculine agreement to other types of targets (for instance,

past participles and predicative adjectives), but never to all instances of gender agreement. As underscored by Wälchli (2017), and contrary to what suggested in previous literature (Rudzīte 1980), the unfolding of these developments varies substantially across different Tamian varieties and cannot be subsumed under one unitary model of change.

For three of the sampled languages, Kundjeyhmi (Central Gunwinyguan), Udi (Lezgic), and Aghul (Lezgic), the patterns of change behind the reduction and loss of gender agreement patterns cannot be fully inferred based on the data at our disposal.

4.5 Reducing gender agreement systems: summary

In our data, the reduction and loss of gender agreement can be described as the result of two distinct processes: morphophonological erosion and redistribution of agreement. We also found evidence for some directional effects in the way in which these developments spread. The morphophonological erosion of gender inflections tends to spread from nouns to those agreement targets that are syntactically more adjacent to nouns (i.e., adnominal modifiers). Conversely, the redistribution of agreement patterns affects anaphoric pronouns (i.e., the agreement targets that are least adjacent to nouns) first. In our sample, these directional effects are attested across different language families and different types of gender systems, which makes it reasonable to hypothesize that they may respond to more general, possibly universal, tendencies in language change. Furthermore, we believe that these directional effects are due to two distinct types of functional constraints: the syntactic cohesion between agreement targets and their controllers, on the one hand, and the sensitivity of agreement targets to semantic properties of discourse referents, on the other hand. The higher the syntactic cohesion (e.g. with definite and indefinite articles), the lower the sensitivity to referential properties, and vice versa (personal pronouns have looser syntactic cohesion with nouns and are therefore more sensitive to semantics). We suggest that the Agreement Hierarchy, a generalization over observed tendencies in the distribution of syntactic and semantic agreement, makes it possible to detect and describe the connection between these two opposite tendencies. This is because, as also outlined in §3, the two ends of the scale, attributive modifiers and personal pronouns, represent instances of highest and lowest degree of syntactic cohesion, and lowest and highest likelihood of semantic agreement, respectively. In §7 we discuss how these different diachronic developments pattern with the evolution of gender complexity.

5 Emerging gender agreement systems

The literature on the rise of grammatical gender is vast, and cannot be reported here in detail. Broadly speaking, two opposite scenarios have been proposed in order to account for the origin of grammatical gender systems. According to the first scenario, the development of classificatory strategies precedes the rise of gender agreement patterns. Gender systems originate from classifiers and classificatory nouns that grammaticalize as agreement markers and, eventually, as gender markers on nouns (Greenberg 1978b; Corbett 1991). According to the second scenario, the development of agreement precedes the development of classificatory distinctions. Nichols (1992: 139–142) argues that the development of classificatory distinctions encroaches on preexisting (person and/or number) agreement patterns whose distribution may be based on covert, in the sense of not morphosyntactically realized, animacy distinctions or on other highly cognitively salient types of distinctions. Against this background, the debate on the origins of grammatical gender systems has focused on a diverse variety of gendered language families, such as Indo-European (Matasović 2004; Luraghi 2011), Atlantic-Congo (Greenberg 1978b; Williamson 1994), Eastern Nilotic (Heine & Vossen 1983), or on individual languages such as the Boran language Miraña (Seifart 2005) or the Southern Daly language Ngan'gityemerri (Reid 1997).

In this section we focus on the hitherto understudied semantic and morphosyntactic properties of young, non-mature (in the sense of Dahl 2004) gender systems. Two main types of young gender agreement systems are brought to attention in this work: (1) emerging gender systems that result from the grammaticalization of light nouns, such as the noun for "woman", as generalized anaphoric devices (see Wälchli 2019 [this volume]) and (2) emerging gender systems that result from the rise of marginal agreement patterns in the domain of adnominal modification, which we discuss in this section. In line with the tendencies also observed for the decline and loss of gender agreement, the two types of emerging gender agreement systems discussed in this volume appear to flag the agreement domains at the two opposite ends of the Agreement Hierarchy (the attributive domain and the anaphoric domain). Neither of these systems, however, originates from classifiers or pre-existing agreement patterns.[19]

While it is impossible to predict whether these emergent patterns of gender agreement will develop into more grammaticalized types of systems, we believe

[19] The emergence of gender agreement from the grammaticalizion of classificatory light nouns is studied, for instance, by Grinevald & Seifart (2004) and Seifart (2005), with a special focus on Amazonian languages.

that they offer a unique insight into the rise of complexity in the domain of gender marking as well as into its stability and transmissibility. In the languages of our sample, the emergence of gender agreement in the domain of adnominal modification can result either from language-internal developments or from language contact. These two cases are discussed separately in the remainder of this section.

5.1 Language-internal development of gender: Nalca

Nalca is a Mek language of the Nuclear Trans-New Guinea family spoken in the Highlands of Tanah Papua. The gender system of Nalca is described by Wälchli (2018), both from a synchronic and diachronic perspective. Nalca has a sex-based gender system, with five gender distinctions and semantic and formal (phonological) assignment; gender distinctions are not overtly coded on nouns and the sole targets of gender agreement are a set of function words, which, beside marking gender, also work as case and deictic marking hosts. The gender markers of Nalca and their respective labels are given in Table 4.

Table 4: Gender in Nalca

Gender	Marker
Masculine (some human males)	*be-*
Feminine (some human females)	*ge-*
Neuter/nouns with Consonant + Vowel phonotactic structure (CV), 'the thing(s) that…'	*ne-*
Default Noun	*e-*
Default Phrase (locative, adverbs)	*a-*

Gender agreement in Nalca is noun phrase internal and strongly tied to linear adjacency between controller nouns and agreement targets. When the adjacency condition is not fulfilled, or when the controller noun is not preceded by attributive adjectives (which favor the expression of gender), inherent gender distinctions are neutralized and the agreement pattern triggered on the case/deictic host is that of the Default Phrase gender *a-*, which is typically used with non-prototypical controllers. This illustrated in (8).

(8) Nalca (Mek; Wälchli 2018: 71)
me: *a-ra* *gelelinga* *sovb-vka*
child(CV) DP-TOP unnoticed enclose.in.netbag-CVB
bo-ba-lam-e:k. *Nauba me:* *ne:-ra* *al-biyvk.* *Me:k*
carry-go-HAB/IPVF-PST.3PL. big child(CV) CV-TOP 3SG-alone. small
me: *ne:-ra* *sovb-vka* *bo-ba-lam-e:k*
child.CV CV-TOP enclose.in.netbag-CVB carry-go-HAB/IPVF-PST.3PL
'They carried the boy away secretly in a netbag. A big boy went by himself. A small boy they carried in a netbag.'

The Nalca noun for 'child' *me:* is Neuter (it has a CV type of phonotactic structure). However neuter agreement is marked only when the noun is accompanied by the attributive modifiers for 'big' and 'small'. When it occurs on its own, as in the first of the three sentences exemplified in (8), the Default Phrase gender agreement *a-* is selected.

Wälchli (2018) describes gender in Nalca as a recent innovation within Mek languages. The gender markers of Nalca have cognates in all related Mek languages, but in none of these languages are these markers part of a system of classificatory distinctions in paradigmatic opposition with each other. In Nalca, an emergent system of nominal classification has resulted from a complex array of multiple, independent patterns of language change. The onset of this evolutionary process is the reinterpretation of a uniqueness/saliency marker targeting the top end of the Animacy Hierarchy (*bi-*) as an agreement marker in opposition with *a-*, probably marking non-uniqueness and low animacy (Wälchli 2018). This type of system is attested in the neighboring languages Eipo and Una, where a high degree of animacy is flagged by the marker *bi-*.

5.2 Contact-induced gender emergence

Contact-induced gender emergence presupposes borrowing of agreement patterns, a phenomenon which is argued to take place only in the context of prolonged contact between two or more speech communities, presupposing child bi-/multilingualism (Thomason & Kaufman 1992; Thomason 2001; Trudgill 2011). The three languages discussed in this section – Chamorro (Austronesian), Lekeitio Basque (Basque), Shumcho (Sino-Tibetan) – fit this scenario in that: (1) they show instances of borrowed gender agreement, (2) they are spoken in a situation of intense and prolonged contact with the languages from which the agreement patterns are borrowed.

We begin our overview of contact-induced gender systems with Chamorro, an independent branch within the Austronesian family, spoken in the Northern Mariana Islands. If borrowed patterns of gender agreement are excluded, in Chamorro, nominal classification is restricted to a small set of classifiers, which are almost exclusively used in possessive constructions. Definite articles vary depending on the information structure status of the nominal they modify (they are sensitive to focus), and there is no gender marking on personal pronouns nor noun-phrase internal agreement, apart from optional multiple plural marking (Stolz 2012: 111). Contact between Chamorro and Spanish starts on an occasional basis during the 16th century, it reaches its apex during the Spanish colonization (between the 17th to end of the 19th century), before it starts declining with the advent of the US occupation, and terminates after World War I. The emergent gender system of Chamorro is described in detail by Stolz (2012). Sex-based gender distinctions manifested through agreement on adnominal modifiers emerged in the language as a result of borrowing of nouns and property words from Spanish. The gender system of Spanish is based on a masculine vs. feminine type of opposition with a combination of semantic, morphological and opaque assignment rules. In Chamorro, the Spanish gender assignment rules are reanalyzed into a predictable system of semantic assignment. Agreement with human female controllers is marked by *-a* (Spanish feminine agreement) while human male controllers, as well as any other type of controller nouns, trigger *-o*/*-u* agreement (Spanish masculine agreement). This is illustrated in example (9).

(9) Chamorro Feminine (a) and Non-Feminine (b) Gender (Austronesian; Stolz 2012: 123)

a. *Ma-nobena-na-ye i mi-milagros-a na Bithen.*
PASS-novena-RED-REF DEF RED-miraculous-F LINK Virgin
'A novena is being conducted for the abundantly miraculous Virgin'.

b. *desde antitites na tiempo esta gof bunit-u na siuda i ya*
since RED:before LINK time already very nice-NF LINK town DEF TN
Hagåtña.
Hagåtña
'A very long time ago, Hagåtña was a very pretty town already'.

In (9b) the Spanish-borrowed noun for town, *siuda*, triggers non-feminine agreement. However its correspondent in Spanish, *ciudad*, is grammatically feminine. Gender assignment in Chamorro is thus predictable based on semantic properties of the controller nouns, and does not fully comply with the assign-

ment rules of the donor language. The Chamorro corpus used by Stolz (2012) reveals 300 pairs of words that are sensitive to the distinction between Feminine and Non-Feminine gender. These can be both property words and nouns. Semantically, they cover a wide range of meanings from physical properties to character traits, from names of professions to kinships, ethnonyms, and young animals with sexual dimorphism (Stolz 2012: 117). Of these gender-sensitive lexical items, the property word *bunitu/a* 'pretty, nice, handsome' is the most frequent token for the encoding of sex-differentiation and agreement. With respect to the productivity of gender marking on nouns, Stolz (2012) finds that Spanish derivational rules for the encoding of gender distinctions on nouns may in some cases extend to Chamorro and English nominal stems as in *dander/a* 'male/female musician' from the Chamorro verb stem *dandan* 'to play music', and in *apostero/a* 'male/female upholsterer' from the English noun *upholsterer*. With respect to the productivity of gender marking outside nouns, adjectival adnominal modifiers borrowed from Spanish may index Feminine Gender when modifying a Chamorro noun denoting a female entity. However, the only set of words that are morphosyntactically suited to mark agreement are adnominal modifiers of Spanish origin. Finally, not all Spanish loanwords are sensitive to gender distinctions and there is a considerable amount of intra-speaker and regional variation as to which words are part of the system of gender distinctions and which are excluded; the range of this variation is still to be studied. In sum, Chamorro displays a semi-productive sex-based type of gender system, where gender assignment is semantically predictable and the only targets of gender agreement are a subset of property words borrowed from Spanish. While the system originated through prolonged and intense contact with Spanish, the evolution of gender agreement in Chamorro grammar and usage continues beyond the disappearance of Spanish as a local contact language, and follows patterns of development that do not completely overlap with those of the donor language.

replaced 'genderless language'

Lekeitio Basque is another example of a language without gender in which marginal patterns of nominal gender marking and gender agreement have intruded through the borrowing of a (small) set of nouns and property words from Spanish, and are used to index semantic properties of discourse referents. Lekeitio Basque is a variety of western Basque spoken in Lekeitio, a town located in the province of Bisqay, within the Spanish Basque Country. According to Hualde et al. (1994: 1–2), Basque is the preferred language of interaction among Lekeitians, even though Lekeitio is a largely bilingual town, with the majority of speakers having an active command of both Basque (standard and local variety) and Spanish. In addition, the authors report that, even though Standard Basque

is the official language of instruction, the local variety is generally preferred to the standard language in everyday communication outside the class environment as well as in formal registers of communication (e.g., communication from the mayor and other local authorities, at church). In Lekeitio Basque, *-a* is used to express reference to female entities, whereas *-o* is used for males. Similarly to the Chamorro case, the borrowed gender suffixes appear both on borrowed nouns, where they qualify as a word formation strategy for the overt coding of natural gender distinctions, and on borrowed modifiers, where they qualify as an instance of gender agreement. Examples of borrowed nouns and modifiers with overt gender distinctions are: *enano/a* 'dwarf'; *álto/a* 'tall'; *alúmno/a* 'student'; *tónto/a* 'stupid, silly', *txúlo/a* 'arrogant' (Hualde et al. 1994: 108–109). Interestingly, gender marking on nouns and adjectives is also extended to Basque lexemes: *gixájo/a* 'poor man/poor woman'; *sorristo/a* 'lousy'; *txotxólo/a* 'stupid, short witted' (Hualde et al. 1994: 109). Finally, when gender-sensitive adjectives are used as a base to derive verbs, gender markers are retained. In such cases, gender is marked through a suffix occurring between the root and the derivational suffix, leading to a pattern of affixation which is unknown to Spanish morphology. This pattern is shown in example (10).[20]

(10) Deadjectival verbs indexing natural gender in Lekeitio Basque (Hualde et al. 1994: 109)
morenotu = 'to become tanned (a male)' < *moréno* 'dark (male)'
morenatu = 'to become tanned (a female)' < *moréna* 'dark (female)'
majotu = 'to become handsome (a male)' < *májo* 'handsome (male)'
majatu = 'to become handsome (a female)' < *mája* 'handsome (male)'

Contact-induced emergence of gender agreement is also attested in the Thebor (Bodic, Sino-Tibetan) language Shumcho, spoken in the Kinnaur district of Himchal Pradesh in the Indian Himalaya, a highly multilingual area at the crossroads between Bodic and Indo-Aryan languages, where Hindi is the language of administration and mass media. In general, natural gender distinctions in Shumcho are encoded lexically; there is no morphological gender marking on nouns and no gender agreement on adjectives and verbs. However, there exist a number of nouns and adjectives for which gender distinctions can be marked suffixally (*-a* = masculine; *-e* = feminine), e.g. *šara/e* 'beautiful', 'young person';

[20] An alternative analysis of the patterns illustrated in (10) is, of course, that the gender-differentiating adjectives are stored as independent lexical items.

laṭa/e 'deaf, dumb', 'deaf/dumb one'.[21] In the majority of cases, these words are of clear Indo-Aryan origin, other cases are less clear. Whenever gender-sensitive adjectives modify nouns denoting humans, gender must be marked, independently of whether the head noun is of Bodic or Indo-Aryan origin (Christian Huber, personal communication). With non-human animates and inanimate nouns gender-sensitive adjectives are invariably feminine. In naturally occurring discourse, however, speakers may sometimes choose to index the biological gender of animals, especially if they feel emotionally attached to them (Christian Huber personal communication; Huber 2011: 76). Some instances of masculine/feminine gender distinctions of the type attested in Shumcho are also found in Jangshung, the other Thebor language included in our sample, as well as in almost all West Himalayish languages; their origin is often connected with loanwords from neighboring Indo-Aryan languages (Christian Huber, personal communication). The distribution and spread of these marginal gender marking systems in the languages of the area are, however, still poorly investigated.

In sum, the three instances of borrowed gender agreement patterns attested in our sample and discussed in this section share a number of characteristics both at the morphosyntactic and semantic level:

1. They result from borrowing of nouns and adjectives, which leads to the emergence of instances of nominal gender marking and of gender agreement patterns, respectively.

2. They are noun-phrase internal.

3. They have purely semantic assignment rules: whatever the gender assignment rules of the donor language, the borrowed agreement patterns are used to signal semantic properties of nouns, and, typically, natural gender distinctions.

Finally, the productivity of these borrowed gender agreement patterns varies a great deal in native speakers' usage and from language to language.

5.3 Emerging gender systems: summary

The number of languages examined in this section is too small to formulate any valid generalization on crosslinguistic properties of young gender systems with

[21] Gendered adjectives can also be used as nouns, in the absence of an overt nominal head (Christian Huber, personal communication).

gender agreement restricted to the domain of adnominal modification. Yet, a couple of remarks can be made on what appear to be recurrent properties of such systems.

Firstly, all four languages examined exhibit non-pervasive gender agreement, which is restricted to one type of target only (case marking hosts in the case of Nalca, borrowed adnominal modifiers in the case of Chamorro, Leiketio Baque, and Shumcho). In all four languages, then, the syntactic cohesion between controllers and targets is maximal, and, in the case of Nalca, also tied to a rather rigid principle of linear adjacency.

Secondly, in all four languages, gender marking is *conditional* rather than *absolute* in the sense that it is constrained by (1) syntactic properties of noun phrases, whereby gender agreement occurs only if the target and the controller noun are adjacent to each other, as in Nalca, or (2) lexical restrictions, whereby only borrowed adjectival modifiers can agree in gender, as in Chamorro, Lekeitio Basque, and Shumcho.

Crosslinguistic similarities between the examined systems are even more striking in the case of contact-induced gender systems. As mentioned before, in the languages examined in this section, emergent gender agreement patterns result from lexical borrowing. Gender marking patterns are transferred along with borrowed nominal and adjectival stems, and the assignment principles that underpin their use in the donor languages are reanalyzed. The resulting assignment systems in the recipient languages are purely semantic in that they especially target the encoding of natural gender distinctions with human (or highly animate) referents. This is suggestive of a possible hierarchical tendency whereby semantic gender assignment rules are preferred to mixed types (semantic and formal) of assignment rules, even if the donor language has both semantic and formal rules. Finally, in the cases examined here, the recipient languages are not genealogically related (apart from Shumcho and Jangshung); they belong to language families that are typically genderless and that, prior to contact, display agreement in other grammatical domains (such as number or person).

It remains to be seen whether the similarities between the three contact-induced emerging gender systems are due to the fact that the donor languages themselves (Spanish, Indo-Aryan languages) have rather homologous, and in fact, genealogically related, gender systems, or whether these similarities speak of more general tendencies with respect to the kind of gender agreement systems that can emerge as a result of language contact (e.g. only semantic, only noun-phrase internal etc.). Only a larger crosslinguistic survey could tackle this question. However, what the instances of contact-induced gender emergence ex-

amined here suggest is that borrowing should be counted as a possible source scenario for the rise of gender systems crosslinguistically.

In §7, we will address how the emergent gender systems surveyed here pattern in terms of complexity.

6 Expanding gender agreement systems

In our sample, the expansion of gender agreement systems is attested under three different scenarios: (1) through the extension of gender marking to new agreement domains via grammaticalization processes (as in the Northwestern Iranian languages Kafteji and Eshtehardi, and in the Khasian languages Pnar and Khasi); (2) as a consequence of contact between languages with different types of gender systems (Michif); and (3) as a result of language planning and standardization (Makanza Lingala). The three scenarios are briefly surveyed in the following.

While the erosion and loss of gender distinctions is not uncommon within Northwestern Iranian varieties (as we observed with the Kelasi case discussed in Section 4.2), in some languages of this group new patterns of gender agreement have grammaticalized in the domain of verbal morphology. In Kafteji, for instance, all tense forms of the intransitive past verb stems inflect for gender in all three singular persons. In Eshtehardi, gender inflection in the domain of verbal morphology is somewhat less pronounced. While intransitive past verbs and copula verbs inflect for gender in the third person singular, only copula verbs inflect for gender even in the first and second person singular. According to Stilo (to appear), the construction through which gender agreement expanded to these domains of verbal inflection is: "PARTICIPLE$^{M/F}$ + COPULA". This construction consisting of participial forms inflecting for gender, followed by copula verb forms, later grammaticalized into a new type of synthetic perfect retaining the gender inflection of the original participial form. The marking of gender distinctions on these recently grammaticalized verb forms is thus directly connected with the source constructions from which these forms originate. The extent to which gender distinctions are marked on verbs across the three person values varies across languages (Stilo to appear: 29).

When compared with each other, the Khasian (Austroasiatic) languages Lyngngam, Pnar and Khasi display a continuum of increasing gender agreement domains. Lyngngam has a pronominal gender system, with gender distinctions marked on personal pronouns and deictic pronominal bases. In Pnar and Khasi, pronominal and deictic markers are used as pre-nominal gender clitics, which mark gender within the noun phrase. In Khasi, the encoding of gender distinc-

tions has also extended to the verbal domain. According to Anne Daladier (personal communication) the pervasiveness of gender agreement and the degree of predictability of assignment rules in these three languages are inversely correlated: the higher the number of agreement targets, the less semantically transparent the gender assignment rules. The distribution of gender agreement systems across the three Khasian languages included in the sample is illustrated in Figure 4. These observations should be tested on a wider set of languages within the family.

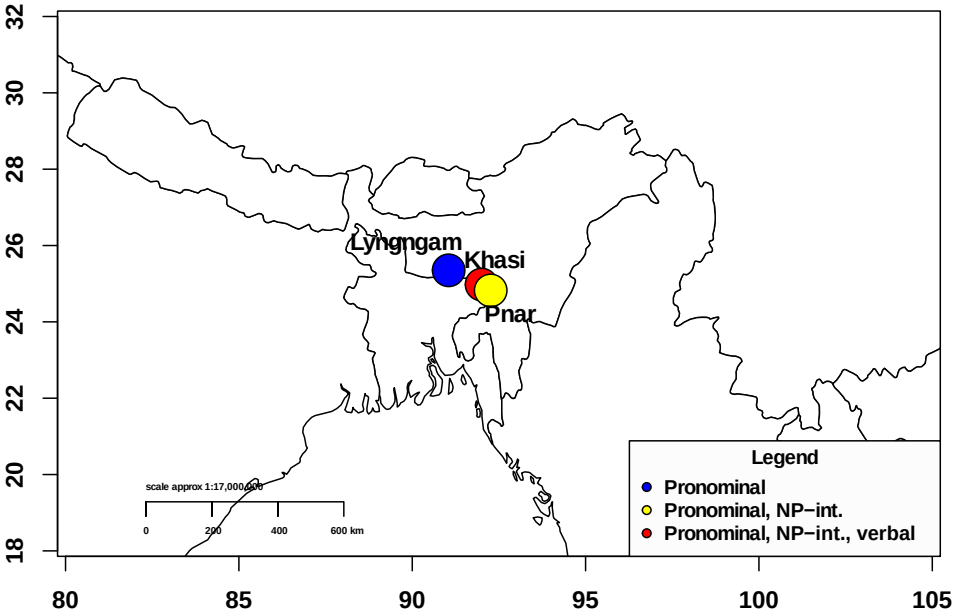

Figure 4: Expansion of gender agreement within Khasian

Michif (scenario 2) is a nearly extinct mixed language originated through intense contact and multilingual practices between female Cree speakers and male French speaking fur trade workers (thoroughly described by Bakker 1997). As a result of these intriguing dynamics of language contact and transmission, the lexicon and morphosyntax of Michif are split into two: nominal lexicon and morphosyntax are French-based while verbal lexicon and morphosyntax are Cree-based. Accordingly, Michif has two co-existing gender systems, with two different systems of gender assignment – sex-based and animacy-based – that manifest themselves through a sharp division between gender agreement within the noun phrase and gender agreement on verbs (with the exception of demonstratives,

which comply to the verb-phrase agreement pattern). The noun-phrase gender system is taken from French, while the verb-phrase gender system is based on Cree. This unique split system of gender agreement is illustrated in (11) where the controller noun for 'mare' triggers feminine agreement within the noun phrase and animate agreement on the verb.

(11) Michif (Mixed Language, Canada and US; Bakker 1997: 87)
la žyma: ki:-aja:w-e:w æ̆ pči pulæ̆
DEF.AN.F.SG mare PST-have-TA.3⟶3^{I} INDEF.AN.M.SG little foal
'The mare had a foal.'

The last instance of expanding gender agreement systems in our sample is Makanza Lingala (scenario 3). In this variety of Lingala, non-sex-based, arbitrary gender distinctions (and corresponding gender agreement patterns) were reintroduced during the standardization process that the language underwent between 1901 and 1902 under the influence of the Scheutist missionaries, who wanted to create an official language that looked more like a 'proper Bantu language'. Kinshasa Lingala, which is nowadays the most widely spoken variety of Lingala and which did not undergo the standardization process attested in Makanza Lingala, exhibits a heavily reduced gender system where gender distinctions and gender agreement patterns are exclusively animacy-based. This reduced gender system is the result of the pidginization and creolization processes that are at the very origins of the history of Lingala, which is the historical descendent of the Bangala pidgin, developed at the Bangala state post on the northwestern banks of the Congo River (for more details on the history of different varieties of Lingala and their gender systems see Bokamba 1977; Di Garbo 2016; Meeuwis 2013).

To summarize, our data suggest that the patterns of change through which languages may acquire more domains of gender inflection tend to be rather heterogeneous and language-specific. However, the limited number of cases examined here does not allow us to formulate any far reaching generalization on the dynamics of gender agreement expansion. While this calls for further investigation, patterns of gender agreement expansion will not be discussed further in the remainder of the paper.

7 How simple/complex are gender agreement systems on the rise and/or in decline?

In §2, we brought up two hypotheses about the complexity of gender systems. Firstly, in viewing the complexity of gender as an evolving variable, instances

of gender systems in decline could be considered as reducing complexity and instances of gender systems on the rise/under expansion as emerging/increasing complexity. Secondly, both reducing and rising gender systems could be expected to show less complexity than their full-fledged counterparts. The data presented in this paper do not, however, support these hypotheses. In this section, we show that many of the processes of reduction and emergence of gender agreement attested in our data contribute to increase the complexity of gender systems as matched against the proposed measures of gender complexity.

Starting with reducing gender agreement, we suggest that especially in those cases in which patterns of reduction only affect sub-parts of the agreement system, whether as a result of morphophonological erosion or of redistribution of agreement, this cannot be described as a straightforward simplification process. In Standard Swedish, for instance, the merger between the Masculine and Feminine genders in the domain of noun-phrase internal agreement gave rise to: (1) a sex-based, referential system of gender assignment, which is active only in the domain of pronominal agreement and for nouns that denote entities at the top end of the animacy hierarchy (humans and, occasionally, higher animals); (2) a non-sex-based, semantic and formal type of gender assignment system, which is active through agreement in the domain of adnominal modification. When mapped onto the model of gender complexity proposed by Audring (2017), this split in the type of classificatory distinctions that agreement targets are sensitive to qualifies as an increase in gender complexity, as illustrated in (12). (The symbol "<" here, as well as in (13), (14) and (15), reads as "less complex than".)

(12) Split agreement system and gender complexity (adapted from Audring 2017)

Matching values (between targets) < Mismatching values (between targets)

This effect can be analyzed as a violation of the Principle of Independence in that the type and number of gender distinctions available in a language vary depending on the type of agreement targets that inflect for gender. Mismatching gender values across different types of targets need to be separately specified in the description of a gender system, which leads to an increase in description length and thus in complexity.

Similarly, we saw that the redistribution of agreement is usually triggered by the reanalysis of the gender assignment of hybrid nouns. In the Asia Minor Greek dialects, for instance, the critical items are nouns that are grammatically masculine or feminine, but semantically denote inanimate entities. In some Asia Minor

Greek varieties (such as Pontic), the ongoing reanalysis of the gender assignment rules associated with these nouns is reflected through mismatching agreement patterns whereby targets adjacent to nouns retain syntactic agreement and non-adjacent targets agree semantically. In Audring's model of gender complexity, hybrid nouns qualify as a "complexifying phenomenon in a gender system" because they engender mismatches in the agreement patterns that they control. This is schematized in (13) and (14).

(13) Hybrid nouns and gender complexity (Audring 2017)

Consistent controller < Hybrid controller

(14) Semantic agreement and gender complexity (Audring 2017)

Targets do not have a choice in value < Targets have a choice in value

When, due to mismatches between grammatical gender and semantic properties of hybrid nouns, agreement targets have a choice in value, these choices need to be specified in the description of a gender system. This increases the description length of the system, and thus its complexity.

Conversely, when the reduction, loss or semantic reanalysis of gender agreement patterns are more pervasive, this usually results in an uncontroversial simplification of the gender agreement system. Under morphophonological erosion, this is for instance the case of English, where sex-based gender distinctions are only preserved on third personal and possessive pronouns and index purely semantic distinctions.[22] Under agreement redistribution, this is the case of Rumeic Greek, where the gender system has become completely semanticized. Nouns denoting male entities are masculine, nouns denoting female entities are feminine, and nouns denoting inanimate entities are neuter.

Moving on to the emergence of gender agreement, the young gender systems examined in this paper also exhibit some features of high complexity when measured against the dimensions proposed by Audring (2017). We observe that, under contact-induced gender emergence, only a subset of lexical items within a given word class (nouns and/or adjectives) is sensitive to gender distinctions. For instance, in Chamorro, only property words borrowed from Spanish can inflect for gender, and there is a great deal of intraspeaker variation as for how productively gender agreement is used. Similarly, in Nalca, where the emergent gender

[22] On the use of the pronouns 'he' and 'she' with inanimate referents in varieties of American and Australian English see Pawley (2004).

system is the result of a language internal development, gender marking is also not fully productive, and it can be switched off whenever certain syntactic conditions within the noun phrase are not met. Low productivity and optionality in gender marking count as complexifying factors according to Audring (2017): they introduce variability in the gender agreement system of a language as a result of lexical and/or grammatical idiosyncrasies that are, in fact, independent of gender.

(15) Low productivity and gender complexity (Audring 2017)

Gender marking is obligatory < Gender marking is optional

Gender marking is fully productive < Only a subset of lexical items per agreement target mark gender

When gender is not fully obligatory or fully productive, specifying explicitly under which circumstances gender marking occurs adds to the system's description length, which means higher complexity. Conversely, the emergent gender systems examined in this paper are rather simple with respect to domains of gender agreement, given that they all display one agreement target, which in all cases examined is confined to the domain of adnominal modification.

Reducing and emerging gender systems represent transitional stages between the *absence of gender* and *full-fledged gender systems*, two rather stable stages in the history of individual languages and language families. These transitional stages are to a large extent associated with phenomena that, we think, increase gender complexity as a side-effect of ongoing language change. In the case of gender reduction, we observed, for instance, a pervasive occurrence of mismatching agreements, which is due to the fact that innovations (a) do not immediately reach all available agreement targets, but rather spread gradually across agreement domains; and (b) do not immediately affect all controller nouns, but rather those with ambiguous semantics (that is, hybrid nouns) first. Under gender emergence, gender agreement tends to be non-obligatory and thus non-frequent. Therefore the main factors underlying increased complexity in reducing and emerging gender systems are partial distributions and optionality, which are ultimately connected to ongoing variation and change.[23] While we hope to have shown that some crosslinguistically recurrent patterns can be associated with these systems in transition, we think that their relative stability is harder to generalize over and depends on the interplay between internal and external dynamics of change, the understanding of which falls outside the scope of this paper.

[23] This has also been pointed out to us by Jenny Audring.

8 Concluding remarks and prospects for future research

We consider the main contribution of this paper to be bringing diachrony in focus in the typological study of gender complexity. We hope to have shown that investigating closely related languages enables us to formulate empirically grounded diachronic inferences about the decline, rise and expansion of gender agreement, as well as about how these dynamics of change affect the complexity of gender systems. In particular, we found that both gender agreement patterns in decline and gender agreement patterns on the rise feature properties of increased complexity when assessed against existing gender complexity metrics. We suggested that emerging and declining patterns of gender agreement represent transitional stages between two poles: genderless languages and full-fledged gender agreement systems. These poles often appear as less complex than the transitional stages, as represented in our sample. Whether this can be generalized over all cases of emerging and declining gender systems is a hypothesis that should be tested on a larger data set and, possibly, with the support of quantitative methodologies.

We think that one additional contribution of this paper is to have shown that implicational hierarchies can be used as schemas for investigating complexity variation across languages in a meaningful way, not only at the synchronic level (as previously suggested by Miestamo 2009), but also diachronically. In this respect, we found that, in the languages of our sample, the agreement domains at the two opposite ends of the Agreement Hierarchy, attributive modifiers and personal pronouns, often function as the place from which processes leading to both the rise and the decline of gender agreement begin. Furthermore, our data suggest that at least the reduction and loss of gender agreement tend to be directional in nature, and that the type of directionality at stake is predicted by whether loss and reduction are due to morphophonological erosion or redistribution of agreement patterns.

We hope that these results may spark further research on the relationship between the complexity of gender systems and other well-known implicational universals in the domain of gender marking, such as the series of implicational universals on the availability of gender distinctions in the plural as opposed to the singular (e.g., Universal 37), or in pronouns as opposed to nouns (e.g., Universal 43), formulated by Greenberg (1963). We believe that this line of research is particularly promising to shed new light on synchronic and diachronic interactions between gender and other grammatical domains, and their effect on the complexity of gender systems.

Finally, one important question that is left out from this paper is whether there are any external factors that contribute to explain why and under which conditions gender agreement systems complexify or simplify. Even though many of the instances of change discussed in this paper clearly involved language contact as a causal factor, the question of the relationship between the evolution of gender agreement systems and language ecology was not addressed systematically here. Thus the answer to this question must be left to further studies. Our impression so far is that gender agreement patterns – whose evolutionary dynamics we have tried to unravel in this paper – might be a better match for the study of the sociolinguistic correlates of gender complexity than, say, sheer number of genders and/or type of assignment systems. Patterns of gender agreement directly hinge on inflectional morphology, which has so far been one of the main foci of research in testing the effects of social structures and language ecologies on the rise and transmissibility of linguistic complexity.

Acknowledgments

We thank two anonymous reviewers, as well as Jenny Audring, Johanna Nichols, and Bernhard Wälchli for providing constructive comments on previous versions of the paper. For financial support, Francesca Di Garbo is thankful to the Wenner-Gren Foundations.

Special abbreviations

The following abbreviations are not found in the Leipzig Glossing Rules:

3⟶3	3rd person animate obviative	LINK	linker
AN	animate	HAB	habitual
CV	consonant + vowel phonotactic structure	NF	non-feminine
		RED	reduplication
DP	default phrase gender	TA	transitive animate verb
INAN	inanimate	TN	toponym
INDEF	indefinite		

References

Agbetsoamedo, Yvonne. 2014. Noun classes in sɛlɛɛ. *Journal of West African Languages* 41(1). 95–124.

Aikhenvald, Alexandra Y. & Robert M. W. Dixon. 1998. Dependencies between grammatical systems. *Language* 74(1). 56–80.
Åkerberg, Bengt. 2012. *Älvdalsk grammatik*. Ulum Dalska.
Audring, Jenny. 2009. *Reinventing pronoun gender*. Vrije Universiteit, Amsterdam. (Doctoral dissertation).
Audring, Jenny. 2014. Gender as a complex feature. *Language Sciences* 43. 5–17.
Audring, Jenny. 2017. Calibrating complexity: How complex is a gender system? *Language Sciences* 60. 53–68.
Audring, Jenny. 2019. Canonical, complex, complicated? In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity*. Vol. 1: *General issues and specific studies*, 15–53. Berlin: Language Science Press.
Bakker, Peter. 1997. *A language of our own: The genesis of Michif, the mixed Cree-French language of the Canadian Métis*. Oxford: Oxford University Press.
Balode, Laimute & Axel Holvoet. 2001. The Latvian language and its dialects. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *The circum-baltic languages: Typology and contact*. Vol. 1: *Past and present*, 3–40. Amsterdam: Benjamins.
Bokamba, Eyamba. 1977. The impact of multilingualism on language structures: The case of central africa. *Anthropological Linguistics* 19(5). 181–202.
Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15(2). 203–224.
Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
Corbett, Greville G. 2010. Implicational hierarchies. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 190–205. Oxford: Oxford University Press.
Curzan, Anne. 2003. *Gender shifts in the history of English*. Cambridge: Cambridge University Press.
Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Barbara Unterbeck (ed.), *Gender in grammar and cognition*. Vol. 1: *Approaches to gender*, 99–115. Berlin: Mouton de Gruyter.
Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: Benjamins.
de Boeck, Egide. 1904. *Grammaire et vocabulaire du lingala, ou langue du Haut-Congo*. Brussels: Polleunis-Ceuterick.

Di Garbo, Francesca. 2014. *Gender and its interaction with number and evaluative morphology: An intra- and intergenealogical typological survey of africa*. Stockholm University. (Doctoral dissertation).

Di Garbo, Francesca. 2015. Questionnaire: Grammatical gender and language ecologies.

Di Garbo, Francesca. 2016. Exploring grammatical complexity crosslinguistically: The case of gender. *Linguistic Discovery* 14(1). 46–85.

Di Garbo, Francesca. forthcoming. The complexity of grammatical gender and language ecology. In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*. Oxford: Oxford University Press.

Dingemanse, Mark. 2009. Noun classification in Siwu. Paper presented at the Conference on African Languages and Linguistics. Leiden 2009.

Duke, Janet. 2010. Gender reduction and loss in Germanic: The Scandinavian, Dutch, and Afrikaans cases studies. In Antje Dammel, Sebastian Kürschner & Damaris Nübling (eds.), *Kontrastive germanistische Linguistik*, 643–672. Hildesheim: Olms.

Evans, Nicholas. 2003. *Bininj Gun-Wok: A pan-dialectal grammar of Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.

Frenda, Alessio. 2011. Gender in Irish between continuity and change. *Folia Linguistica* 45(2). 283–316.

Gblem-Poidi, Massanvi Honorine. 2007. Nominal classes and concord in Igo (Ahlon). In Mary Esther Kropp Dakubu, George Akanlig-Pare, Kweku E. Osam & Kofi K. Saah (eds.), *Proceedings of the annual colloquium of the Legon-Trondheim Linguistics Project 10–20 January 2005*, vol. 4, 52–60. Legon: Linguistics Department, University of Ghana.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Greenberg, Joseph H. 1978a. Diachrony, synchrony and language universals. In Joseph H. Greenberg (ed.), *Universals of human language*. Vol. 1: *Method and theory*, 61–92. Stanford: Stanford University Press.

Greenberg, Joseph H. 1978b. How does a language acquire gender markers? In Joseph H. Greenberg, Charles Ferguson & Edith Moravcisk (eds.), *Universals of human language*. Vol. 3: *Word structure*, 47–82. Stanford: Stanford University Press.

Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages: Towards a comparison. *Linguistic Typology* 8(2). 243–285.

Güldemann, Tom & Ines Fiedler. 2019. Niger-Congo "noun classes" conflate gender with declension. In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity*. Vol. 1: *General issues and specific studies*, 97–148. Berlin: Language Science Press.

Hammarström, Harald, Robert Forkel & Martin Haspelmath (eds.). 2018. *Glottolog 3.2*. Jena: Max Planck Institute for the Science of Human History. http://glottolog.org/.

Heine, Bernd & Rainer Vossen. 1983. On the origin of gender in Eastern Nilotic. In Rainer Vossen & Marianne Becchaus-Gerst (eds.), *Nilotic studies: Proceedings of the international symposium on languages and history of the nilotic peoples, cologne, january 4-6, 1982*, 245–268. Cologne: Dietrich Reimer Verlag.

Hualde, José Ignacio, Gorka Elordieta & Arantzazu Elordeta. 1994. *The Basque dialect of Lekeitio*. Bilbo: Universidad del País Vasco/Euskal Herriko Univertsitatea.

Hualde, José Ignacio & Jon Ortiz de Urbina. 2003. *A grammar of Basque*. Berlin: Mouton de Gruyter.

Huber, Christian. 2011. Some notes on gender and number marking in Shumcho. In Gerda Lechleitner & Christian Liebl (eds.), *Jahrbuch des Phonogrammarchivs*, vol. 2, 52–90. Göttingen: Cuvillier Verlag.

Huldén, Lars. 1972. Genussystemet i Karleby och Nedervetil. *Folkmålsstudier* 22. 47–82.

Hultman, Oskar Fredrik. 1894. *De östsvenska dialekterna*. Helsinki: Svenska landsmålsföreningen.

Karatsareas, Petros. 2009. The loss of grammatical gender in Cappadocian Greek. *Transactions of the Philological Society* 107. 196–230.

Karatsareas, Petros. 2014. On the diachrony of gender in Asia Minor Greek: The development of semantic agreement in Pontic. *Language Sciences* 43. 77–101.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The circum-baltic languages: An areal-typological approach. In Östen Dahl & Maria Koptjevskaja-Tamm (eds.), *Circum-baltic languages: Typology and contact*, vol. 2, 615–750. Amsterdam: Benjamins.

Kusters, Wouter. 2003. *Linguistic complexity: The influence of social change on verbal inflections*. University of Leiden. (Doctoral dissertation). Utrecht: LOT.

Leufkens, Sterre. 2015. *Transparency in language: a typological study*. Amesterdam: University of Amsterdam. (Doctoral dissertation).

Luraghi, Silvia. 2011. The origin of the proto-indo-european gender system: Typological considerations. *Folia Linguistica* 45(2). 435–464.

Maho, Jouni. 1999. *A comparative study of bantu noun classes*. Vol. 13 (Orientalia et Africana Gothoburgensia). Göteborg: Acta universitatis gothoburgensis.

Matasović, Ranko. 2004. *Gender in indo-european*. Heidelberg: Winter.

Meeuwis, Michael. 2013. Lingala. In Susanne Michaelis, Philipe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages*. Vol. 3: *Contact languages based on languages from Africa, Asia, Australia and the Americas*, 25–33. Oxford: Oxford University Press.

Miestamo, Matti. 2006. On the feasibility of complexity metrics. In Krista Kerge & Maria-Maren Sepper (eds.), *Finest Linguistics. Proceedings of the Annual Finnish and Estonian Conference of Linguistics, Tallin, May 6–7, 2004*, 11–26. Tallin: TLÜ.

Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 23–41. Amsterdam: Benjamins.

Miestamo, Matti. 2009. Implicational hierarchies and grammatical complexity. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*. 80-97: Oxford University Press.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Nichols, Johanna. 2003. Diversity and stability in language. In Brian D. Joseph & Richard D. Janda (eds.), *The handbook of historical linguistics*, 283–310. Oxford: Blackwell.

Passer, Matthias Benjamin. 2016. *The typology and diachrony of nominal classification*. University of Amsterdam. (Doctoral dissertation).

Pawley, Andrew. 2004. Using *he* and *she* for inanimate referents in English: Questions of grammar and world view. In Nick J. Enfield (ed.), *Ethnosyntax, explorations in grammar and culture*, 110–137. Oxford: Oxford University Press.

Priestly, Tom M. S. 1983. On 'drift' in Indo-European gender systems. *The Journal of Indo-European Studies* 11(3-4). 339–363.

Reid, Nicholas. 1997. Class and classifier in ngan'gityemerri. In Mark Harvey & Nicholas Reid (eds.), *Nominal classification in aboriginal australia*, 165–225. Amsterdam: Benjamins.

Rudzīte, Marta. 1980. Läripärast liivi morfoologias. In *Congressus quintus internationalis fenno-ugristarum. Turku, 20-27. VIII 1980*, vol. Pars VI, 231–236.

Sasse, Hans-Jürgen. 1993. Syntactic categories and subcategories. In Joachim Jakobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: Ein internationales Handbuch zeitgenössicher Forschung/ An international handbook of contemporary research*, 646–686. Berlin: Walter de Gruter.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña*. Radboud Universiteit Nijmegen. (Doctoral dissertation).

Seifart, Frank. 2010. Nominal classification. *Language and Linguistics Compass* 4(8). 719–736.

Siemund, Peter & Florian Dolberg. 2011. From lexical to referential gender: An analysis of gender change in medieval English based on two historical documents. *Folia Linguistica* 45(2). 489–534. DOI 10.1515/flin.2011.018.

Soubrier, Aude. 2013. *Description de l'ikposso uwi*. Lyon: Université Lumière Lyon 2. (Doctoral dissertation).

Stilo, Donald. to appear. Loss vs. expansion of gender in Tatic languages: Kafteji (Kabatei) and Kelasi.

Stolz, Thomas. 2012. Survival in a niche. On gender-copy in Chamorro (and sundry languages). In Martine Vanhove, Thomas Stolz, Hitomi Otsuka & Aina Urdtze (eds.), *Morphologies in contact*, 93–140. Munich: Akademie-Verlag.

Svärd, Erik. 2019. Gender in New Guinea. In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity*. Vol. 1: *General issues and specific studies*, 227–280. Berlin: Language Science Press.

Thomason, Sarah G. 2001. *Language contact: an introduction*. Washington D.C.: Georgetown University Press.

Thomason, Sarah G. 2015. When is the diffusion of inflectional morphology not dispreferred? In Francesco Gardani, Peter Arkadiev & Nino Amiridze (eds.), *Borrowed morphology*, 27–46. Berlin: Mouton de Gruyter.

Thomason, Sarah G. & Terrence Kaufman. 1992. *Language contact, creolization and genetic linguistics*. Berkeley & Los Angeles: University of California Press.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. New York: Oxford University Press.

Wälchli, Bernhard. 2017. The incomplete story of feminine gender loss in Northwestern Latvian dialects. *Baltic Linguistics* 8. 143–214.

Wälchli, Bernhard. 2018. The rise of gender in Nalca (Mek, Tanah Papua): The drift towards the canonical gender attractor. In Sebastian Fedden, Jenny Audring & Greville G. Corbett (eds.), *Non-canonical gender systems*, 68–99. Oxford: Oxford University Press.

Wälchli, Bernhard. 2019. The feminine anaphoric gender gram, incipient gender marking, maturity, and extracting anaphoric gender markers from parallel texts. In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity*. Vol. 2: *World-wide comparative studies*, 61–132. Berlin: Language Science Press.

Williamson, Kay. 1994. Niger-Congo overview. In John Bendor-Samuel (ed.), *The Niger-Congo languages*, 3–45. Lanham, New York, London: University Press of America.

Yarshater, Ehsan. 1969. *A grammar of Southern Tati dialects*. The Hague - Paris: Mouton.

Appendix

The sampled genealogical units are listed by macroarea and higher levels of classification are mentioned, if applicable. (Q) indicates that, for any particular language, data have been collected through full questionnaire responses; (p.c.) stands for personal communication (i.e., data collected through consultation of language experts but no full questionnaire response).

Family by macroarea	Language	Glottocode	Source
AFRICA			
Bantu (Atlantic-Congo)			
	Kinshasa Lingala	ling1263	Bokamba (1977); Meeuwis (2013)
	Makanza Lingala	ling1269	de Boeck (1904); Bokamba (1977); Meeuwis (2013)
Ghana-Togo-Mountain (Atlantic-Congo)			
	Selee	sele1249	Agbetsoamedo (2014)
	Igo	igoo1238	Gblem-Poidi (2007; p.c.)
	Ikposo	ikpo1238	Soubrier (2013); Ines Fiedler (p.c.)
AUSTRALIA			
Gunwinggu (Central Gunwinyguan, Gunwinyguan)			
	Kunwinjku	gunw1252	Evans (2003)
	Kundjeyhmi	gunw1252	Evans (2003)
	Kune	gunw1252	Evans (2003)
EURASIA			
Khasian (Austroasiatic)			
	Khasi	khas1269	Anne Daladier (p.c.)
	Lyngngam	lyng1241	Anne Daladier (p.c.)
	Pnar	pnar1238	Anne Daladier (p.c.)
Basque			
	Standard Basque	basq1248	Hualde & Ortiz de Urbina (2003)
	Lekeitio Basque	bisc1236	Hualde et al. (1994)
Balto-Slavic (Indo-European)			
	Latvian	latv1249	Balode & Holvoet (2001), Anna Kalnaca (p.c.)
	Tamian Latvian	latv1249	Balode & Holvoet (2001); Thomason (2015); Koptjevskaja-Tamm & Wälchli (2001)
Greek (Indo-European)			
	Modern Greek	mode1248	Karatsareas (2009; 2014)
	Pontic Greek	pont1253	Karatsareas (2009; 2014) (Q)
	Rumeic Greek	mari1411	Karatsareas (2009; 2014)
	Cappadocian Greek	capp1239	Karatsareas (2009; 2014)

Continued

Family by macroarea	Language	Glottocode	Source
EURASIA			
Insular Celtic (Indo-European)			
	Irish	iris1253	Frenda (2011)
	Irish (Ros Much)	conn1243	Frenda (2011)
North Germanic (Indo-European)			
	Elfdalian	dic (ISO)	Åkerberg (2012); Östen Dahl (Q)
	Karleby Swedish	oste1241	Hultman (1894); Huldén (1972)
	Standard Swedish	swed1254	Duke (2010); Mikael Parkvall (Q)
Northwestern Iranian (Indo-European)			
	Eshtehardi	esht1238	Stilo (to appear); Yarshater (1969)
	Kafteji	kaba1276	Stilo (to appear; p.c.)
	Kelasi	kaba1276	Stilo (to appear; p.c.)
Lezgic (Nakh-Daghestanian)			
	Archi	arch1244	Michael Daniel, Nina Dobrushina (Q)
	Aghul	aghu1253	Nina Dobrushina (Q)
	Udi	udii1243	Nichols (2003);Wolfgang Schulze (Q)
Thebor (Bodic, Tibeto-Burman)			
	Shumcho	shum1243	Huber (2011; p.c.)
	Jangshung	jang1254	Huber (2011; p.c.)
NORTH AMERICA			
Mixed Language			
	Michif	mich1243	Bakker (1997)
PAPUNESIA			
Chamorro (Austronesian)			
	Chamorro	cham1312	Stolz (2012)
Mek (Nuclear-Trans-New-Guinea)			
	Nalca	nalc1240	Wälchli (2018)
	Eipo	eipo1242	Wälchli (2018)

Name index

Language index

Grammatical gender and linguistic complexity

The many facets of grammatical gender remain one of the most fruitful areas of linguistic research, and pose fascinating questions about the origins and development of complexity in language. The present work is a two-volume collection of 13 chapters (plus an introductory chapter in each volume) on the topic of grammatical gender seen through the prism of linguistic complexity. The contributions discuss what counts as complex and/or simple in grammatical gender systems and whether the distribution of gender systems across the world's languages relates to the language ecology and social history of speech communities. Contributors demonstrate how the complexity of gender systems can be studied synchronically, both in individual languages and over large cross-linguistic samples, and diachronically, by exploring how gender systems change over time. Volume two consists of three chapters providing diachronic and typological case studies, followed by a final chapter discussing old and new theoretical and empirical challenges in the study of the dynamics of gender complexity.

This volume is preceded by volume I: *General issues and specific studies*, which, in addition to three chapters on the theoretical foundations of gender complexity, contains six chapters on grammatical gender and complexity in individual languages and language families of Africa, New Guinea, and South Asia.

ISBN 978-3-96110-180-1