Chapter 4

Why is gender so complex? Some typological considerations

Johanna Nichols

University of California, Berkeley
Higher School of Economics, Moscow
University of Helsinki

A cross-linguistic survey shows that languages with gender can have very high levels of morphological complexity, especially where gender is coexponential with case as in many Indo-European languages. If languages with gender are complex overall, apart from their gender, then gender can be regarded as an epiphenomenon of overall language complexity that tends to arise only as an incidental complication in already complex morphological systems. I test and falsify that hypothesis; apart from the gender paradigms themselves, gender languages are no more complex than others. The same is shown for the other main classificatory categories of nouns, numeral classifiers and possessive classes. Person, the other important indexation category, proves to be less complex, and I propose that the reason for this is that person, but not gender, is referential, allowing hierarchical patterning to emerge as a decomplexifying mechanism.

Keywords: gender, case, numeral classifiers, possessive classes, person hierarchy, referential, inflection, canonical complexity, simplification, diachronic stability.

1 Introduction

There can be little doubt that gender systems are complex, and in various ways: compare the large number of gender classes in Bantu languages, the intricate and opaque fusion with case, number, and declension class in conservative Indo-European languages, the extensive allomorphy of Tsakhur gender agreement

(Nakh-Daghestanian; examples below), or the semantically unpredictable genders of Spanish or French nouns. Even for Avar (Nakh-Daghestanian), which has a three-gender system with almost no allomorphy of gender markers and complete semantic predictability, there is a random division of verbs into those that take gender agreement and those that do not. The open question about the complexity of gender systems is why? Here I propose an answer based on two factors: one is the inexorable growth of complexity as a maturation phenomenon that can continue indefinitely unless braked by some simplification process (Dahl 2004; Trudgill 2011), and the other is a self-correcting measure that is available to some agreement categories but not to gender, for reasons probably having to do with referentiality.

Two different ways of measuring and comparing complexity will be used here. The first is what I will call *inventory complexity*, which goes by various names (e.g. Dahl 2004: *resources*, Miestamo 2008: *taxonomic complexity*, Di Garbo & Miestamo 2019 [in Volume II]: *the principle of fewer distinctions*): the number of elements in the inventory or values in a system, for some domain such as the number of phonemes, tones, genders, classifiers, derivation types, basic alignments, or basic word orders, or the degree of verb inflectional synthesis. Inventory complexity figures in Dahl (2004), Shosted (2006), Nichols (2009), Donohue & Nichols (2011), and many other works. It is not a very accurate or satisfactory measure of complexity, not least because it does not measure non-transparency, which is the kind of complexity that has been shown to be shaped by sociolinguistics (Trudgill 2011); but it is straightforward to calculate (though data gathering can be laborious), and appears to correlate reasonably well with other, better measures of complexity. Below I use inventory complexity to compare complexity levels of different languages for the practical reason that there is an existing database of inventory complexity (that of Nichols 2009, subsequently expanded) which counts items across several phonological, morphological, and syntactic subsystems across 200 languages.

The other measure used here is *descriptive complexity* or *Kolmogorov complexity*: the amount of information required to describe a system. This is a better measure and captures well the non-transparency relevant to learnability and prone to be shaped by sociolinguistics, but it is very difficult to measure and compare. Here I follow Nichols (2016; forthcoming) in using canonicality theory (Corbett 2007; 2013; 2015; and others) as an approximate measure of descriptive complexity (though not an exact equivalent; some differences are noted below); see Audring (2017) for a similar approach. Canonicality theory is not primarily a

complexity measure but a theoretical undertaking that aims at improving definitions and technical understanding of linguistic notions. It defines a logical space (for a linguistic concept or structure or system) by determining the central, or ideal, position in that space and attested kinds of departures from that ideal, and measuring non-canonicality as the extent of departure (or number of departures) from the ideal. A central notion in defining the ideal position is the structuralist notion of biuniqueness, or one form, one function; any departure from that ideal is non-canonical. The literature on canonicality offers a good deal of work on morphological paradigms, which makes it a straightforward matter to count the number of non-canonicalities in a paradigm. I use canonicality theory partly because of the availability of this previous work and partly because it is well grounded in morphological theory (and taken seriously by theoreticians) yet applicable on its own without requiring adoption of an entire comprehensive formal framework. I survey this kind of complexity with a different database that samples morphological subsystems as sparingly as possible in order to keep the survey manageable (underway; 80 languages so far).

In what follows I illustrate descriptive complexity with some inflectional paradigms and show how much information grammars need to present (and do present) to adequately describe some of those paradigms (§2); this shows that the presence of gender in a paradigm can make it extremely complex by the inventory metric. But is it the gender morphology itself that is complex? Or is gender rather an epiphenomenon of overall language complexity, a category that tends to arise only as an incidental complication in already complex morphological systems? §3 and §4 raise and falsify the hypothesis that gender – and classification more generally – is embedded primarily in already complex languages, showing that it is gender itself that is complex. §5 compares the complexity levels of person, the other important indexation category. It appears that descriptive complexity easily becomes great in the indexation categories, and that person has recourse to self-correcting, self-simplifying mechanisms that gender lacks. More precisely, person has means of self-correction and self-simplification other than sheer reduction of inventory size or overall loss of the category – apparently unlike gender. This partly accounts for the great diachronic stability of gender systems (Matasović 2014) and in particular the remarkable stability of complexity in gender systems. The reason for the different behavior of gender and person appears to be that person, but not gender, is referential. The concluding section (§6) considers some ramifications of this claim.

2 Complexity in gender: Examples and measurement

Gender systems can be complex in themselves and also in the way that they interact with other inflectional categories. This section compares some more and less complex gender systems and proposes a way to quantify their complexity. Examples come from the database of non-canonicality, which samples small but easily comparable inflectional subsystems from a few basic parts of grammar in order to get some view of complexity across the inflectional system: marking of A, S, O, G, T, and possessor roles on nouns; the same forms of inflectional pronouns; singular A and O marking in the most basic past and nonpast synthetic forms of verbs; inflectional classes of affixes for nouns, pronouns, and verbs; and inflectional classes of stems for all three.

The paradigms in Tables 1–2 show the inflection of nouns in four grammatical cases in the singular of Mongolian (which has no gender) and Russian (which has three genders).

Table 1: Mongolian (Khalkha; Svantesson 2003: 163, Janhunen 2012: 297–298, 106–112, 66–68; Janhunen's transcription). Extension underlined.

	'book'	'year'
Nominative	*nom*	*or*
Genitive	*nom-ÿn*	*or-n̲-ÿ*
Accusative	*nom-ÿg*	*or-ÿg*
Dative	*nom-d*	*oro-n̲-d*

Table 2: Russian (M = masculine, F = feminine, N = neuter). Extension underlined.

	'brother' M.anim.	'house' M.inan.	'book' F	'window' N	'net' Fourth	'time' Fourth, Extended
Nom.	*brat*	*dom*	*knig-a*	*okn-o*	*set'*	*vremja*
Gen.	*brat-a*	*dom-a*	*knig-i*	*okn-a*	*set-i*	*vrem-en̲-i*
Acc.	*brat-a*	*dom*	*knig-u*	*okn-o*	*set'*	*vremja*
Dat.	*brat-u*	*dom-u*	*knig-e*	*okn-u*	*set-i*	*vrem-en̲-i*

Mongolian has only one declension class in terms of suffixes. There are some differences in suffixes (not shown), all predictable from the phonology of the stem (its final consonant and vowel harmony class). There are two stem classes: simple nouns as in 'book', and one with an *-n-* extension in certain cases, as in 'year'. In Russian matters are more complex. There are four declension classes of suffixes: those of 'brother' and 'house', 'book', 'window', and 'net' and 'time' in Table 2, plus a class of indeclinables not shown.[1] There is a minor class of stems with extensions, illustrated here with the *-en-* extension of 'time'. The animate and inanimate masculine nouns differ in their accusative allomorphs; they are largely predictable from the animacy of the referent. Further subclasses not shown here are mostly phonological and predictable from the final consonant or stress position of the stem. (Plural forms and the other oblique cases, not part of this survey, would add further non-canonicalities.)

In canonicality theory, declension classes are non-canonical because they contribute nothing; the one-form-one-function ideal is violated because a declension class has form but no function. There are two kinds of inflectional classes: those involving stems and those involving the inflectional affixes (Bickel & Nichols 2007: 184). Traditionally recognized inflectional classes may be based on stems, affixes, or both, but I factor these out here. A stem declension class has stem change or extension which is a form without meaning; a declension class of affixes is a set of forms but the set has no meaning. The canonical situation is to have no declension classes, so Mongolian is canonical as to affixes (and nearly so as to stems) but Russian is not. On the other hand, if there are declension classes, then they should all be different, since the point of declension classes is differentiation. Affix classes should have affixes all of which are different from the affixes of other classes; each stem class should have an extension, ablaut, stress shift, or whatever that is unique to it. Here Russian declension is non-canonical because there are a number of syncretisms between classes, e.g. the *-u* dative of masculine and neuter declensions or the *-i* genitive of feminine and fourth declensions. Furthermore, within declension classes case affixes should all be different from each other, with one affix per case. Here Russian declension is non-canonical because

---

[1]For this breakdown of the Russian declension classes see Corbett (1982). The traditional terminology deals only with declension classes of endings and not with stem classes. The first three classes are now, at least in work in English, commonly called masculine, feminine, and neuter for the noun genders prototypically or exclusively associated with their members: masculines are only masculine, feminines mostly feminine, neuters only neuter. There is no standard synchronic term for the class of 'net' and 'time'; I call it the fourth declension. Traditionally, the masculine and neuter classes have been grouped together for historical reasons: both go back to the Indo-European o-stem declension. The traditional terms are first declension (masculine and neuter), second (feminine), and third ('net' and 'time').

there are many syncretisms within paradigms, such as genitive and accusative for masculine animates or genitive and dative in 'net' and 'time' in Table 2. A different departure from the principle of a single affix per case is the allomorphy of the accusative ending in the masculine declension: -*a* for animates but zero for inanimates. This is a split of one category into two forms, sensitive to some additional category.[2] (For the general claims of canonicality theory in this paragraph see Corbett 2007; 2013; 2015.)

Thus, of the forms surveyed here, while Mongolian case inflection has one morphological non-canonicality in the system, Russian has 11: the intra-paradigm syncretisms of masculine animate genitive-accusative, inanimate nominative-accusative, neuter nominative-accusative, fourth declension nominative-accusative and genitive-dative; the -*en*- extension in 'time'; the allomorphy of suffixes between animate and inanimate masculines; and the inter-paradigm syncretisms of nominative zero suffix (masculine and fourth), genitive -*a* (masculine, neuter), genitive -*i* (feminine, fourth), and dative -*u* (masculine, neuter).[3] Both languages have further non-canonicalities in parts of their noun inflectional paradigms that are not surveyed here.

The common types of non-canonicalities in inflectional paradigms are listed in Table 3. All depart from the ideal of one form, one function.

The complexity measurements for the Mongolian and Russian systems shown above are given in Table 4 and Table 5. They pertain only to singular declension; in Mongolian the plural adds no more non-canonicalities, as in the separative morphology of the language plural and case are marked by different morphemes (and the case suffixes are largely the same as in the singular), while in Russian plurality and case are coexponential, with a single suffix signaling the two categories.

Thus a descriptively and theoretically adequate synchronic grammar of Mongolian needs to display only two paradigms, while for Russian five must be shown.

---

[2]Whether there is a category of animacy that these case forms signal, mark, etc. or they are sensitive to animacy but do not carry it as a category meaning is a thorny issue that cannot be solved here. I will speak of sensitivity to a category (or indeed a property that is not necessarily an actual category of the language) without taking a stance on the larger issue.

[3]Since the extensions of Mongolian appear in some but not all non-nominative cases, perhaps that distribution should also be counted as a non-canonicality, giving Mongolian a total of two. The non-predictability of the Mongolian extension is greater than for Russian: it appears in some but not all non-nominative cases, while the Russian one can be analyzed as appearing in all non-nominative cases (with that pattern then overlain by the nominative-accusative syncretism, which gives an unextended stem to the accusative as well). It is, incidentally, coincidence that the extension has the same consonant in the two languages and appears in the same cases of the partial paradigms shown in Table 1 and Table 2.

Table 3: Non-canonicalities in inflectional paradigms, and their numbers of forms and functions. 2 (+): two or more. 0*: perhaps defectivity involves not a zero function but an actual function that is blocked from realization.

	Forms	Functions
Syncretism	1	2 (+)
Zero affixes	0	1
Fused exponence (coexponence) of categories	1	2 (+)
Allomorphy, splits	2 (+)	1
Defectivity (gaps)	0	0*

Table 4: Inventory complexity for Mongolian and Russian singular core grammatical cases

	Declensions	Genders
Mongolian	1	0
Russian	5	3, plus animacy

Table 5: Descriptive complexity for Mongolian and Russian singular core grammatical cases. The phonological information is the description in the phonology of automatic alternations.

Mongolian noun paradigms	Russian noun paradigms
Display 1 paradigm, plus 1 extended	Display 5 paradigms, plus extended (2 extension allomorphs)
Access phonological information	Access phonological information Comment on syncretisms, allomorphy, etc.

Pedagogical grammars will usually display more, and, at least for Russian, automatic phonological and morphophonological alternations involving plain vs. palatalized stem-final consonants trigger orthographic changes and are usually also included in the paradigm display. I will not attempt to measure the amount of information presented in the commentaries, notes, etc. on declension paradigms in the two languages, but at first glance it appears to be no less extensive per declension class for Russian than for Mongolian. In any event the difference of one vs. five paradigms suffices to show that more information is required for describing noun declension in Russian than Mongolian.

Russian declension is more complex than Mongolian declension because late Proto-Slavic fused into single case suffixes what had been a sequence of separate stem-forming suffixes (essentially, extensions) plus what had been a more uniform set of case endings in late Proto-Indo-European. The IE extensions had some correlation with gender, and this has tended to increase over time in the attested daughter languages, spurred in no small part by the fact that gender agreement was signalled in adjectives by shifting back and forth between what were lexical or word-formation categories for nouns: *o*-stem suffixes were used for masculine and neuter agreement, the *a*-stem suffixes for feminine. This means that the fusion of gender into the case-number paradigms, an accident of Proto-Slavic sound changes, received support in the gender agreement paradigms of adjectives. This seems to have stabilized the system despite the non-transparency introduced by adding gender to the mix.

Now consider what makes for complexity in a gender system with no fusion of categories or markers. Table 6 shows the gender class markers for Ingush, a Nakh-Daghestanian language of the central Caucasus. Every noun belongs to a gender (usually covert on the noun) marked by root-initial agreement on some verbs and adjectives. Nouns and pronouns referring to male humans belong to V gender, females to J gender; this is what I will call referent-based gender assignment,[4] where gender is predictable from (in this case) the sex of the referent. In the plural both take B agreement, except that first and second person pronouns take D in the plural.[5] Other nouns are arbitrarily assigned to one or another of B, J, and D gender. Altogether there are eight gender classes consisting of singular-plural pairs, and four gender markers. The gender markers have no allomorphy (other than the split of singular B gender into D and B plurals, for which allomorphy is one possible analysis) and no fusion with other segments or morphemes, and are

---

[4]This is the *referential gender* of Dahl (2000). I use *referential* in a different sense; see note 14 below.

[5]In recent linguistic work on Nakh languages the genders are named for the letter name of their marker.

Table 6: Ingush gender markers (Nichols 2011: 144)

	Singular	Plural	Examples
1st, 2nd person pronouns	*v/j*	*d*	me, you, us
3rd person pronouns (human)	*v/j*	*b*	him, her, them
male human nouns	*v*	*b*	man, Ahmed
female human nouns	*j*	*b*	woman, Easet
some animals, inanimates	*b*	*d*	ox, head
some plants, inanimates	*b*	*b*	apple, family
inanimates, some animals	*j*	*j*	wolf, fence
inanimates, some animals	*d*	*d*	dog, house

Table 7: Gender agreement in two Ingush verbs. A dot segments off the gender marker. Verbs shown in the simple present tense. (D gender is the citation form.)

	d.ouz- 'know' (kennen)	*dwa=chy=d.uoda* 'go down'
V	*v.oudz*	*dwa=chy=v.uoda*
J	*j.oudz*	*dwa=chy=j.uoda*
B	*b.oudz*	*dwa=chy=b.uoda*
D	*d.oudz*	*dwa=chy=d.uoda*

thus formally transparent. Semantically, as in nearly all gender systems, gender is transparently predictable (referent-based) for nouns and pronouns referring to humans but arbitrary, i.e. opaque, for others.

Formal simplicity vs. complexity is illustrated by the verb paradigms for Ingush and Tsakhur (another Nakh-Daghestanian language: Daghestanian branch, Lezgian subbranch) in Table 7 and Table 8. In Ingush the system is quite transparent: there is no allomorphy and no allophony of gender markers; gender agreement is always root-initial (and the proclitics in Table 7 are readily identifiable from their prosody, some of their segmental phonology, and the fact that they are separable, occurring in word-final positions when the verb is in second position). In Tsakhur it is quite opaque. There is a good deal of allomorphy, and this produces different patterns of syncretism: genders 1 and 4 syncretize in 'hold' but 1 and 2 in 'hang'.[6] Gender is partly prefixal and partly infixal: infixal in formerly bipartite stems, where a former prefix has entrapped the root-initial gender

---

[6]In recent linguistic work on Daghestanian languages the genders are arbitrarily numbered.

marker, but the bipartite structure is ancient and not synchronically transparent. In both languages some but not all verbs take gender agreement: about 30% in Ingush and a very large majority in Tsakhur. Whether a verb takes agreement or not is then highly predictable for Tsakhur but much less predictable for Ingush; in this regard Ingush is less canonical.

In Tsakhur as in Ingush, the first two genders are used of humans and are referent-based, and the last two are arbitrarily assigned. In Avar (Nakh-Daghestanian; Daghestanian branch, Avar-Andic-Tsezic subbranch), gender is formally even simpler than in Ingush (in that for Avar there are no other verb prefixes and no proclitics, so gender markers are not just root-initial but word-initial) and entirely referent-based (there are three genders: masculine, feminine, and other, a.k.a. neuter). Also, unlike Ingush, the plural gender marker is entirely predictable from the singular one. The system is smaller than that of Ingush: three genders and four gender markers for Avar vs. eight genders and four markers for Ingush. The sole non-canonicality of Avar is that not all verbs and not all adjectives take gender agreement (about half of the verbs do, thus unpredictability is maximal).[7]

To summarize this section, non-canonicality can be a good guide to complexity and makes it possible to compare relative degrees of complexity using existing and straightforward criteria. Russian noun declension is considerably more complex than Mongolian; Tsakhur gender agreement is considerably more complex than that of Ingush or Avar; Ingush gender agreement is somewhat more complex than that of Avar. I have not attempted here a calculation of absolute complexity levels based on canonicality. (For a more detailed discussion of non-canonicality as complexity measure see Nichols 2016; forthcoming.)

---

[7]Avar is known for rampant multiple agreement in phrases and clauses: not only verbs and adjectives but also a number of adverbs, determiners, and other forms take agreement (Kibrik 1985; Kibrik 2003). There are three possible analyses of multiple agreement in canonicality theory: (1) Gender is unnecessary, hence non-canonical in itself, so minimizing its use is canonical. (2) Multiple agreement is neutral, as long as all targets receive the same feature values (Corbett & Fedden 2016: 513) and agreement is obligatory (Corbett 2006: 14–15). (3) Given that gender exists, multiple agreement is canonical in that it demonstrates exhaustiveness of features across lexical classes (Corbett 2013: 54) and functional in that it increases consistency and identifiability of gender across different constituents and different utterances. I have no stance on this, but the sociolinguistic history of Avar may be relevant, as Avar is a spreading and inter-ethnic contact language of the type expected to undergo simplification (Trudgill 2011). In contrast, Ingush has undergone a poorly understood spread but is not an inter-ethnic or contact language, and Tsakhur is a small highland language and sociolinguistically quite isolated in Trudgill's sense (in which sociolinguistic isolation means no history of absorbing adult L2 learners; Tsakhur, like other highland Daghestanian languages, has very few adult L2 learners but is not at all isolated from contact of other kinds). If the spreading and inter-ethnic language has extensive multiple agreement, it may well be functional in some way, though canonicality and functionality are different things and not expected to coincide.

Table 8: Gender agreement in two Tsakhur verbs. Aorist tense. (Dobrushina 1999: 85 with some retranscription. *qq* = geminate, *y* = high back unrounded vowel, *X* = uvular.) Dot in citation form marks insertion point and boundary between the gender marker and the pieces of a bipartite stem. In actual inflected forms the gender marker has a dot on either side.

	a.q- 'hold'	*giwa.X-* 'hang'
1	*a.q.qy*	*giwa.r.Xyn*
2	*a.j.qy*	*giwa.r.Xyn*
3	*a.w.qu*	*giwa.p.Xyn*
4	*a.q.qy*	*giwa.t.Xyn*

3 Are gender languages more complex overall?

A possible explanation for the evolution of gender is that it arises easily, as some kind of excrescence or emergent category and probably due to reanalysis of existing markers, in a language that is already morphologically complex and already has at least some agreement as a model for gender agreement. And indeed, gender is almost never the sole inflectional category, or even just the sole agreement category.[8] If gender presupposes complexity, the synchronic result should be that when gender is disregarded languages with gender should still have higher overall complexity than languages without gender. To determine that, this section tests three hypotheses about the overall complexity of languages with and without gender. For all three I use the inventory complexity database of Nichols (2009), expanded to 196 languages with reasonably diverse genealogical and geographical distribution. It should be cautioned, though, that the database is slanted toward inflectional morphology of indexation and head marking, with better representation of categories such as person and classification than e.g. case or other categories of non-heads.[9]

---

[8] A possible exception is the western Nakh-Daghestanian languages, including Ingush and Avar discussed here, where there is no person agreement at all, but only gender agreement. (Arguably there is also number agreement, though that is usually treated as it is in Bantu languages, with number just a matter of gender pairing between singular and plural classes.)

[9] The reason for the imbalance is historical: the morphological measures are mostly drawn from the Autotyp database (Bickel et al. 2017), for which data on NP structure and noun inflection is a more recent addition and still incomplete. This is one reason why the database is best viewed as a convenience sample of categories than as a balanced sample of categories (much less an accurate measure of overall morphological complexity or even just overall complexity of inflectional morphology).

Hypothesis (i): Languages with gender are more complex overall than those without gender. For this count I used the entire set of complexity measures (phonological, morphological, syntactic, lexical), excluding gender; that is, measuring complexity other than in gender. The results are shown in Table 9: there is no significant difference in complexity between gender languages and genderless languages. What little correlation does show up is negative, contradicting the hypothesis.

Table 9: Overall complexity of languages with and without gender.

	Above	Below mean complexity
Gender	28	38
No gender	58	78	n.s. (p = 0.18; Fisher 1-tailed)

Hypothesis (ii): Gender languages are more complex morphologically than genderless languages. This test uses the same survey except that only the morphological measures of complexity are counted. There is a significant negative correlation; see Table 10. Hypothesis (ii) fails, as does the null hypothesis; the finding here is that gender languages are less complex morphologically than genderless languages.[10]

Table 10: Overall morphological complexity of languages with and without gender. Figures in bold are above the expected values.

	Above	Below mean complexity
Gender	15	**43**
No gender	**60**	76	(p = 0.01; Fisher 1-tailed)

Hypothesis (iii): Gender languages have higher inflectional synthesis of the verb than genderless languages. Verb inflectional synthesis was defined as Categories per word (including roles) following the Autotyp database (Bickel et al. 2017). Again the hypothesis is falsified (Table 11).[11]

---

[10]But recall again the bias toward features of heads in the database, above in the text and note 9; to evaluate the impact of Table 10 it is especially important to have a balanced survey of categories.

[11]What small correlation emerges is negative. Bickel & Nichols (2013a) exclude role marking from verb synthesis; on that measure, there is a significant negative correlation, falsifying both survey and null hypotheses and suggesting that it is non-complexity that favors gender. Again (see notes 9 and 10) the result shows that a balanced morphological survey is important.

Table 11: Overall inflectional synthesis of the verb for languages with and without gender.

	Above	Below mean complexity	
Gender	22	36	
No gender	64	75	n.s. (p = 0.19, Fisher 1-tailed)

Thus, except for gender itself, on three criteria gender languages are no more complex than others and may even be less complex. The rise of gender must be due to something other than sheer complexity, and the synchrony of gender does not require or favor overall high complexity.

4 Complexity in classifier systems: numeral classifiers, possessive classification

Perhaps systems of classification in general are complex, so that complexity is not just a peculiarity of gender. This section considers the complexity of numeral classifier and possessive classifier systems.

Numeral classifiers are well known from many East Asian languages, e.g. Mandarin. The systems tend to be large (50 or more in common use for Mandarin, plus many more that can be extracted from occasional occurrence in the long and varied written tradition of Chinese); the inventory complexity is therefore high. The numeral classifiers generally have independent phonological wordhood status and minimal or no sandhi, fusion, etc. and are semantically transparent, though with some flexibility as to what nouns take what classifiers (the flexibility is itself semantically motivated); therefore the descriptive complexity is low.

Elsewhere around the Pacific Rim numeral classifiers tend to be less transparent. Nivkh (isolate; Sakhalin Island and the lower Amur, eastern Siberia) has some 30 numeral classes (Mattissen 2003 gives the highest number) (moderate-high inventory complexity), in which the classifier is fused to the numeral, the combination being semi-transparent, and (at least in the recent and present situation of speech-community contraction and reduced functionality) different classifiers have different distributions: some classifiers apply only to the numerals 1–5, some to 1–5 and 10, and some to all of 1–10 (this is fairly high descriptive complexity). Yurok (Algic, northern California; Robins 1958: 86–91) has 15 classes (moderate inventory complexity), semantically motivated (human, plant, various shapes, etc.). The classifier is inextricably and opaquely fused with the numeral,

yielding a de facto system of 15 classes of numerals (high descriptive complexity). ("Several informants were aware of this complexity and would say admiringly of another speaker that he or she 'knows the numbers' or 'can count in Indian' ": Robins 1958: 87n.)[12] The languages with numeral classifiers range from morphologically non-complex (Mandarin and other Southeast Asian languages) to morphologically complex (Yurok), with the major hotbed of numeral classifier systems found in the morphologically relatively simple languages of Southeast Asia but other languages with numeral classifiers sprinkled all around the Pacific Rim, where languages have high complexity in general. A preliminary conclusion is that numeral classifier systems can be complex in themselves but numeral classifier languages as a set are not more complex than others.

Possessive classes (Nichols & Bickel 2013; Bickel & Nichols 2013b) involve covert classification of nouns which becomes overt only when the noun has possessive morphology. Many languages have a distinction of two classes of nouns, usually termed alienable and inalienable. The formal difference can be as simple as obligatory possession of inalienables vs. optional possession of alienables, and the semantic opposition can be quite straightforward (e.g. kin terms and/or body parts vs. other nouns). In such a language (the most frequent type), both inventory and descriptive complexity are low. A complex system is that of Anêm (isolate, New Britain; Thurston 1982), in which possessed nouns fall into at least 20 classes marked by some simple and some composite suffixes and involving a mix of partly semantic and partly arbitrary classification (Thurston 1982: 37–38), very high inventory complexity. There is a good deal of syncretism between classes, and class membership is semantically unpredictable, so descriptive complexity is also high. The most complex system I have observed is that of Cayuvava (isolate, Bolivia; Key 1967), in which possessive morphemes are circumfixes with much allomorphy of both pieces and partial interdependence between the pieces. Both prefixal and suffixal parts appear to reflect person, and the suffixal part is also purely classificatory. The choice of classifier is semantically unpredictable. The set of first person singular forms is shown in Table 12. The inventory complexity is high and the descriptive complexity might be described as stratospheric.

Thus possessive classification, like numeral classification, can also be quite complex, and probably no less complex than gender. The overall complexity of

---

[12] Mattissen (2003) compiled the fullest list of Nivkh numeral classifiers by cross-tabulating lower figures reported in other sources. Robins compiled his list in analogous fashion from different speakers ("The table…was compiled from several informants and represents a collation of material from all of them, each accepting, though not necessarily volunteering, all the forms tabulated" [87]).

Table 12: First person singular possessive circumfixes in Cayuvava (Key 1967).

a-	…	*-i*
		-ro
		-Ø
		-ai
i-	…	*-i*
		-Ø
ub-	…	*-i*
ku-	…	*-i*
či- ~ ič-	…	*-i*
č-	…	*-ri*

languages with possessive classification ranges from low (as in Polynesian languages: see e.g. Wilson 1982 for Polynesian possessive classification) to high (e.g. Anêm, whose Austronesian-speaking neighbors consider it impossible to learn; Thurston 1982: 51).

Results of the same kinds of tests, for morphological complexity against presence vs. absence of numeral classifiers, possessive classes, or either one are shown in Tables 13–15. Again none of the results are significant: languages with classification of either type are not more complex than those without. There is, however, an interesting trend for a positive correlation of possessive classification and high complexity (Table 14), which merits testing on a larger sample.

Overall, then, neither gender, numeral classifiers, nor possessive classification appears to require or favor general morphological complexity as a diachronic prerequisite or synchronic correlate, and complex classification is not just a simple reflection of the overall complexity level of the language.

Table 13: Overall morphological complexity of languages with and without numeral classifiers

	Above	Below mean complexity	
Classifiers	14	16	
No classifiers	88	80	n.s. (p = 0.35; Fisher 1-tailed)

Table 14: Overall morphological complexity of languages with and without possessive classification

	Above	Below mean complexity	
Poss. classes	38	45	
No poss. classes	41	74	n.s. (p = 0.099; Fisher 1-tailed)

Table 15: Overall morphological complexity of languages with and without classification (numeral or possessive)

	Above	Below mean complexity	
Classification	41	45	
No classification	33	50	n.s. (p = 0.19; Fisher 1-tailed)

5 Complexity in person indexation

Person, like gender, is primarily an agreement or indexation category, and in fact person is the clausal agreement category par excellence. Person indexation on verbs can be quite complex, and this section compares complexity and the evolution of complexity or non-complexity in gender and person systems, arguing that complex person marking systems can develop emergent alternative analyses that are simpler while gender systems do not and apparently cannot do this.

Inventory complexity of person marking is high in West Caucasian languages such as Adyghe and Abkhaz, which index six person-number categories for three roles, for an 18-cell total paradigm; Yimas (Lower Sepik-Ramu, New Guinea; Foley 1991) with 3 persons × 3 numbers × 2 roles (also 18), or Kiowa (Kiowa-Tanoan, U.S.; Watkins & McKenzie 1984), 3 persons × 3 numbers × 2 roles × 2 conjugation classes, plus direct/inverse marking for 17 subject-object paradigm cells (total of 53). In the West Caucasian languages transparency is high, since each argument is indexed by an unambiguous person-number marker in a separate slot, while transparency for Kiowa is low, since subject and object roles are indexed with mostly fused morphemes (see the paradigms in Watkins & McKenzie 1984: 115–116). The Kiowa non-transparencies and the two conjugation classes are non-canonical.

A different kind of non-canonicality is found in languages such as Laz (Kartvelian, Georgia and Turkey; Lacroix 2009: 283, Öztürk & Pöchtrager 2011: 48),

Table 16: Arhavi Laz subject and object indexation paradigm. Only one argument is overt. ... = root + thematic suffix. Phonological alternations not shown. (Lacroix 2009: 283, 298, plus examples on other pages; s.a. Öztürk & Pöchtrager 2011: 51.)

	S/A-	O- ...	-S/A	Examples	
1SG	*b-*	*m-*		*b-dzir-om*	'I see him'
				m-dzir-om	'you$_{sg}$ see me'
2SG		*g-*		*dzir-om*	'you$_{sg}$ see him'
				g-dzir-om	
3SG			*-s/n/u*	*dzir-om-s*	'he sees him'
				m-dzir-om-s	'he sees me'
1PL	*b-*			*b-dzir-om-t*	'we see him'
2PL		*g-*		*dzir-om-t*	'you$_{pl}$ see him'
3PL			*-an/nan/es/n*	*dzir-om-an*	'they see him'

where the two arguments of transitive verbs compete for a single person prefix slot and the competition is resolved by person and role hierarchies (1, 2 > 3, A > O). See Table 16, especially the first two forms listed, where the prefix is first person singular, subject in the first example *b-dzirom* and object in the second *m-dzirom*. The system is non-canonical in that the same slot can mark either subject or object, and in that second person has no overt marking at all. In addition to person/number prefixes, number is also indicated by a plural affix that registers plurality of any argument (A, S, O, G) if it is first or second person, and another that indexes number for a third person S/A.[13] This is non-canonical in that a single category (plural) is marked with different formatives that have different distributions (third person subject indexation vs. non-third-person plural argument registration).

The argument indexation system of Tundra Yukagir (isolate, Siberia; Maslova 2003b) is even less canonical; see Table 17. The system is a proximate/obviative one somewhat like those of Tagalog, Algonquian languages, and others (see Bickel

---

[13]I use *index* and *register* as in Nichols (1992: 48–49): indexation copies or otherwise marks features of the argument (person, number, etc.) on the verb, while registration simply indicates the presence of an argument in the clause but does not agree with or copy features. I assume that what is called promiscuous number marking (Leer 1991) is not indexation (of number on an argument marker, because the argument is not specified) but registration (of a multiple argument, a category similar to pluractionality and easily overlapping with it: see Wood 2007, Yu 2003).

Table 17: Tundra Yukagir obviation system (Maslova 2003b: 18). Focus = proximate. S focus column constructed from other tables in Maslova (2003b) and Kolyma Yukagir (Maslova 2003a).

	Neutral transitive	O focus	A focus intransitive	Neutral	S focus
1sg	*-Ø-ng*	*-me-ng*	*-Ø*	*-je-ng*	*-l*
2sg	*-me-k*	*-me-ng*	*-Ø*	*-je-k*	*-l*
3sg	*-m-Ø*	*-me-le*	*-Ø*	*-j-Ø*	*-l*
1pl	*-j*	*-l*	*-Ø*	*-je-l'i*	*-l*
2pl	*-mk*	*-mk*	*-Ø*	*-je-mut*	*-l*
3pl	*-nga*	*-ngu-me-le*	*-ngu-Ø*	*-ngi*	*-ngu-l*

2011 for this typology), in which one of the arguments is designated as proximate (usually because of topicality or a similar parameter) and the others are obviative. Verb indexation and noun case track proximate and obviative status. (The term for 'proximate' in Tagalog and Yukagir descriptions is usually *focus*.) In Yukagir, unlike other languages with obviation, a proximate argument is not required, and unlike Tagalog the proximate argument can be only A, S, or O (for Kolyma Yukagir, only A or S; Maslova 2003a). Identifying single-function forms that index person/number categories is impossible for most of the cells. Nearly every cell in Table 17 exhibits one or more non-canonicalities.

To judge from the languages surveyed here, person systems can have greater inventory complexity and greater descriptive complexity (more non-canonicalities) than gender systems. However, person systems also have simpler and more canonical analyses available than gender systems do: hierarchical structuring, in which different patterns that violate biuniqueness reduce to a single ordering principle. The Laz paradigm shown in Table 16 reduces to a set of signs plus two hierarchical patterns: 1, 2 > 3 and A > O (for discussion of the Pazar Laz hierarchies see Öztürk & Pöchtrager 2011: 48). Maslova (2003b: 17, 20) reduces much of the complexity and non-transparency of Table 17 to the two hierarchies illustrated in Tables 18 and 19 and summarized in Table 20.

On this perspective, the Yukagir system is still less than straightforward, and it differs from better-known obviation systems in that it tracks the proximate/obviative status of the O while the others mainly track the A. But the individual morphemes are better motivated and the whole system emerges as less non-canonical than the non-hierarchical one, and thus as less complex.

Table 18: Tundra Yukagir obviation: Distribution of transitive markers (Maslova 2003b: 17). Bracketed comment mine.

Person of A:	A focus	Neutral	O focus
1	*-Ø-*	*-Ø-*	*-me-*
1+ other [*i.e.* 1PL]	*-Ø-*	*-j*	*-l*
Non-1	*-Ø-*		*-m(e)-*

Hierarchy: Focus > Speaker > other. Zero suffix signals that A outranks O in this hierarchy.

Table 19: Tundra Yukagir obviation: Person slot (the second element of the internally hyphenated forms in Table 17) in the O focus paradigm (Maslova 2003b: 20)

	O neutral	O focus	
1SG	*-ng*	*-ng*	A = SAP
2SG	*-k*	*-ng*	A = SAP
2PL	*-k*	*-k*	A = 2 + 3 [*i.e.* 2PL]

Hierarchy: SAP > other

Table 20: Summary of hierarchical effects in Tundra Yukagir obviation. (Recall that Focus = proximate.)

Hierarchy	What it determines
Obviation: Focus > Speaker > other	Form of person/number markers
Role: A > O	Zero vs. nonzero suffix
Person: SAP > other	Form of second slot in person/number marker

All forms index the A (relying on hierarchies) and register an O.
Hierarchy for access to O registration: Focus > all else.

A striking example comes from Alutor (Chukchi-Kamchatkan). Paradigms, too long to reproduce here, for the most basic forms are in Nagayama (2003), Mal'ce-va (1998), and others; full tables are in Kibrik et al. (2004: 639–648). The tables are not only long but complex and with dauntingly little correlation of form to function, either within or across paradigms. Kibrik (2003) reduces the forms to a basic person hierarchy of 1SG, 1PL, 2SG > 2PL, 3 for access to the A slot, the reverse for access to O, for relatively polar A and O (and additional provisions for less polar A and O), plus different cutoffs in different mood categories based in part on the speaker's control over, or ability to predict, the event.

Hierarchically based indexation (in which I also include inverse indexation) has the advantage that less information is required than for standard paradigm-based accounts. Roles and/or person can be inferred from hierarchies rather than being fully specified. Those hierarchies are not part of the description of each paradigm; they are grammar-wide, to some extent even universal, as are cross-linguistically favored cutoff points such as 1, 2 > 3 person or S/A > O. For purposes of assessing descriptive complexity, a grammar-wide principle does not have to be specified for particular paradigms and adds no information to their description; a universal principle does not contribute information to any particular grammar.

In these respects, hierarchical indexation may well be canonical. Viewed in the proper perspective, it is not a type of paradigm but what might be called a BLUEPRINT for creating paradigms and forms. Henceforth I will use the term BLUEPRINT because it is not a precise theoretical term and because it implies an instruction or algorithm or the like rather than a structure or set of forms. (How to implement hierarchical and other blueprints in theoretical morphology is a challenge not addressed here.) The paradigm is the blueprint's output, and available evidence indicates that describing the output requires more information than describing the blueprint.

A cross-linguistically recurrent minimal hierarchical system shows up in verbs indexing two arguments, where combinations of first and second person ('I VERB you', 'you VERB me') are often opaque, or overtly mark only one of the persons, or are ambiguous or otherwise non-transparent (Heath 1991; 1998). This amounts to treating the participant scenario not as a pair of arguments and not even as a morphologically fused dyad but as a monad. From what is left unarticulated, plus culture-specific and universal expectations, one can infer who does what to whom; see Heath's detailed analysis. This too is a type of blueprint.

The theoretical claim of Kibrik (2003: 376) for Alutor is that identical forms point to proximity in cognitive space, and the structure of that space is much

less complex than traditional conjugation tables. This statement, and other descriptions of hierarchies, strike me as presenting a view of an alternate, simpler paradigm, but nonetheless a paradigm and not a blueprint.

Person differs from gender and other agreement and classification categories in that only person exhibits hierarchical patterning. Gender and classifiers never do, in my experience. Even in the concurrent gender and classifier system of Mian described by Corbett & Fedden (2016), where one might expect the two systems to compete for a single slot at least in some circumstances, this does not happen. Number and gender can of course be drawn into the patterning of person if they are drawn along in coexponential markers, but on their own they do not form hierarchies.

The reason for this may lie in the fact that person markers are typically, perhaps always, referential. There are three views on whether person markers are referential. One view is that person markers are always referential, not only the pronominal arguments of pro-drop languages but also the person agreement affixes of languages like English or German or Russian, where there is generally a clearly referential overt argument as well as the verbal person marker whose referentiality is at issue (Kibrik 2011). The second view is that person markers are never referential, even in pro-drop languages, but reference arises from the context and the arguments and is attributed to markers in processing or grammatical analysis (Evans 1999; 2003). The third view is that some person markers are referential and some are not: those variously described as pronominal arguments or cross-reference are referential while those described as agreement are not referential but are simply categories of referring NPs (Hengeveld 2012). Whichever view one adopts, it is probably safe to say that if anything is referential in verb indexation, person is. That is, in proneness to referentiality, person > other categories.

I doubt that categories other than person are ever referential. Gender, in particular, appears to never be referential.[14] Creissels (2014) shows that verbs in Avar (Nakh-Daghestanian, eastern Caucasus) are entirely ambiguous between anaphoric, unspecified, and absent readings of one or more arguments. (1) gives examples parallel to his from Ingush, where the grammar is identical in this respect. Ingush can be described as having two zero pronominals, one anaphoric

---

[14] I use *referential* of gender in the same way as I used it of person in the previous paragraph, so that *is referential* means 'refers' or 'can refer'. This is the usage of Kibrik (2011). It is not to be confused with the same word in Dahl's distinction (2000) of referential gender (= my referent-based gender) vs. lexical gender. Both senses of the word are established in the literature; I chose the one having to do with a new point made here, though Dahl's term is probably the earlier one. The issue needs to be resolved; my *referent-based* is only a patch.

and one unspecified, and the first two readings have these as A argument. The third reading has no A at all; this kind of clause, in which the A is absent but the O remains an O and is not promoted to S, is not found as a major clause type in European languages.[15] (2) shows that exactly the same readings are available to a verb that does not take gender agreement (recall from above that gender is a partial category in Ingush). This shows that gender has nothing to do with referentiality in Ingush. (No argument can be made for either Ingush or Avar about referentiality of person, as both languages lack an inflectional category of person.)

(1) Ingush[16]

  a. Anaphoric zero:
      Ø yz   v.iira
      X_i 3sg V.killed
      '(I/you/he/she/they) killed him.'

  b. Unspecified zero:
      Ø   yz   v.iira
      unsp 3sg V.killed

      He was killed (by someone);
      (Someone) killed him;
      'They killed him.'

  c. Absent A:
      yz   v.iira
      3sg V.killed

      'He was/got killed.'

(2) Ingush

  a. Anaphoric zero:
      Ø yz   leacar
      X_i 3sg V.caught

      '(I/you/he/she/they) caught him.'

---

[15]It is not that this verb has ambitransitive (labile) valence; in Ingush this construction seems to be available to all transitive verbs and perhaps all two-argument verbs more generally. Actual ambitransitive valence of the type (A)O occurs in very few Ingush verbs (I know of only the five listed in Nichols 2011: 466–467).

[16]All verbs in (1)–(2) are in the witnessed past tense (a.k.a. aorist). The nonwitnessed tense (*v.iina.v*, *leacaa.v*), which is resultative and/or inferential evidential, would probably be more likely for the (c) examples.

b. Unspecified zero:

Ø   *yz   leacar*
UNSP 3sɢ V.caught

'He was caught (by someone)';
'(Someone) caught him';
'They caught him.'

c. Absent A:

yz   leacar
3sɢ V.caught

'He was caught/arrested.'

All reviewers of this chapter, and most audiences where I have presented this part of it, raise the objection that gender is referential: it is referential in English pronouns, and gender is known to be important in reference tracking. The point merits a brief excursus. As background, saying that a morpheme or category is referential means that it refers, or carries reference, or bears a referential index. If a category is referential, the category itself is what refers, and not the word that carries that category. English pronouns certainly refer, but it is the pronoun and not its gender that is referential. English pronouns are no more (and no less) referential than those of e.g. Finnish or Turkish (languages which have no gender in either nouns or pronouns) or Ingush (which has noun gender but no pronoun gender), or for that matter French or Russian (which have gender in nouns and pronouns). The presence or absence of gender in pronouns, or whether the gender (in languages that have it) is entirely natural (as in English) or agrees with a noun antecedent (as in French or Russian), does not affect the referentiality of pronouns.

Gender has indeed often been said to be useful in reference tracking, but in fact its usefulness in this function is marginal, as human protagonists of narrative and discourse often belong to the same gender. Kibrik (2011: 334–360) makes this claim and supports it with cross-linguistic, discourse, and experimental evidence, and also emphasizes that reference tracking is not the same as referring: reference tracking mostly has to do with disambiguating and resolving potential referential conflicts.

To summarize on referentiality, person can be referential, and perhaps person is always, and necessarily, referential; but gender is not referential.[17] Numeral

---

[17]My own strong intuition is that inflected verb forms in Ingush do not refer. A context like the anaphoric one in (1a)–(2a) can make it unambiguous who performed and underwent the action, and the choice of witnessed vs. non-witnessed evidentiality categories can make clear whether the speaker knows who did what, but the verb form itself does not refer and the gender at most guides the search for an antecedent by narrowing down its possible gender.

classifiers and possessive classifiers are probably also not referential, but as they appear in NPs rather than on verbs the question of referentiality is less clear.[18]

I am not aware that the matter has been the subject of research, but I suggest a diachronic scenario like the following. On verbs that index two arguments, and especially when person agreement develops enough complexity and/or opacity (e.g. in fusion of forms), hierarchical patterns can arise. The most likely first step occurs when phonological change has made formerly discrete A and O person markers opaque and universal person hierarchies step in to disambiguate, and in doing so they impose their own order. Hierarchical structure is thus an emergent pattern, and it functions not in the usual way that paradigms and sets of forms do but in a new way, as a blueprint. A blueprint is functional where complexity is high, because it reduces the complexity. The ability to function referentially seems to be critical to this emergence, perhaps because referentiality makes it possible to draw on universal hierarchies and fix 1<>2 person forms as morphologically opaque monads.[19]

The reason why gender systems can be so complex is then that they have no self-correcting mechanism like the hierarchical blueprint that might simplify them, and they are stable enough that complexity can build up over time without causing the whole system to be shed. Not only are they stable within families; the complex interaction of gender with case and number persisted in Latin, ancient Greek, late Proto-Slavic, and early Germanic, despite large spreads with absorption of substantial numbers of L2 learners, circumstances that are expected to simplify languages but did not appreciably simplify the paradigms of these languages.

The papers by Liljegren (2019 [this volume]) and Di Garbo & Miestamo (2019 [in Volume II]) (and also Maho 1999) show examples of gender systems simplifying, but the way in which they go about simplifying supports my point. Both papers describe changes in which closer alignment of semantics and gender classification occurs in individual words, beginning with a few words and at the extreme ends of Corbett's agreement hierarchy (1991: 248–259). Typically, a word referring to an animate or human but with an arbitrary gender classification begins to trigger an appropriate animate or human gender agreement marker in limited contexts (such as predicate nominal). Over time, more words and more contexts are involved, and eventually the system ends up based on animacy rather than on

---

[18]Numeral classifiers can fuse to demonstratives and those can be referential and can furthermore be accreted to verbs as indexes, but by that point they have begun to function as third person markers which also index classificatory categories.

[19]1<>2 is Heath's now widely used notation for opaque morphemes that are ambiguously 1>2 and 2>1 (1998).

arbitrary classifications. The early stages, however, add complexity, as the gender agreement rules refer to contexts, create alternations and options for some words but not others, and otherwise introduce variation. Alternatively, gender can be lost when gender agreement is lost, and in the languages Di Garbo and Miestamo study, where singular and plural nouns mostly have different gender agreement markers and gender is marked not only in agreement but also on the nouns themselves, the former gender marking changes into a system of number marking. But these are all developments where gender is ultimately simplified by reduction or loss, while I am talking about complex person systems which retain all their categories and markers but in some kind of reanalysis acquire an emergent alternative analysis as blueprint-driven. For this, I believe, we have no analog in gender.

6 Stability of gender

Gender is very stable in language families (Matasović 2007; 2014). In Indo-European, gender – the categories, the markers, and the complex interaction with case paradigms – lasts as long as the original case endings do, so the original system is still largely in place in Baltic and Slavic and to some extent in Germanic (where parts of it are recognizable to the specialist). More precisely, gender does not outlast the original case endings – nor, usually, vice versa (though Armenian is a counterexample: see Kulikov 2006). Even when case was lost in the various Romance languages and in Macedonian and Bulgarian, the gender categories have remained and their markers continue those of early Indo-European. Whatever the reason for this stability, it means that a gender system can evolve considerable complexity without much risk that the language will abandon it or restructure it. The complexity of the Slavic gender system is simplified not by restructuring but by losing case entirely, in Macedonian and Bulgarian; this removes all the complexity that is due to cumulative expression of case with gender, discussed in §2 above. In general in Indo-European, where gender has been lost, case has generally also been lost, as in English or some Iranian languages (e.g. Persian). Loss of gender has happened in three languages and one additional dialect of Nakh-Daghestanian, a very old family (probably older than Indo-European) with about 40 daughter languages, so 10% or less of the family has lost gender. In these languages gender is not cumulative with case but is expressed only in agreement, and languages that lose gender keep case. The languages that have lost gender have histories of large spreads and contact of the kind expected to simplify languages; but not all of the languages with similar sociolinguistic histories have lost gender. The prehistory of gender in Nakh-Daghestanian is still

poorly understood (though see Schulze 1998), but the complexity of gender marking in Tsakhur, discussed above, is a clearly secondary phenomenon caused by positional sound changes after the accretion of spatial prefixes entrapped the gender prefixes. Some high-contact languages have reduced the number of gender markers and categories, but gender is retained and the agreement rules function in much the same way across the family.

Neither the inventory and descriptive complexity of Nakh-Daghestanian gender, nor the descriptive complexity of conservative Indo-European languages, nor any other gender system I am aware of, has any self-correcting mechanism like hierarchical patterning for person.

Acknowledgments

References

Audring, Jenny. 2017. Calibrating complexity: How complex is a gender system? *Language Sciences* 60. 53–68.

Bickel, Balthasar. 2011. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 399–444. Oxford: Oxford University Press.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, 2nd edn., vol. 3, 169–240. Cambridge: Cambridge University Press.

Bickel, Balthasar & Johanna Nichols. 2013a. Inflectional synthesis of the verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/22.

Bickel, Balthasar & Johanna Nichols. 2013b. Obligatory possessive inflection. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/58.

Bickel, Balthasar, Johanna Nichols, Taras Zakharko, Alena Witzlack-Makarevich, Kristine Hildebrandt, Michael Rießler, Lennart Bierkandt, Fernando Zúñiga & John B. Lowe. 2017. *The AUTOTYP typological databases.* https://github.com/autotyp/autotyp-data/tree/0.1.0. Version 0.1.0.

Corbett, Greville G. 1982. Gender in Russian: An account of gender specification and its relation to declension. *Russian Linguistics* 2. 197–232.

Corbett, Greville G. 1991. *Gender.* Cambridge: Cambridge University Press.

Corbett, Greville G. 2006. *Agreement.* Cambridge: Cambridge University Press.

Corbett, Greville G. 2007. Canonical typology, suppletion, and possible words. *Language* 83(1). 8–42.

Corbett, Greville G. 2013. Canonical morphosyntactic features. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 48–65. Oxford: Oxford University Press.

Corbett, Greville G. 2015. Morphosyntactic complexity: A typology of lexical splits. *Language* 91(1). 145–193.

Corbett, Greville G. & Sebastian Fedden. 2016. Canonical gender. *Journal of Linguistics* 52(3). 495–531.

Creissels, Denis. 2014. P-lability and radical P-alignment. *Linguistics* 52(4). 911–944.

Dahl, Östen. 2000. Animacy and the notion of semantic gender. In Barbara Unterbeck (ed.), *Gender in grammar and cognition.* Vol. 1: *Animacy and the notion of semantic gender: Approaches to gender*, 99–115. Berlin: Mouton de Gruyter.

Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity.* Amsterdam: John Benjamins.

Di Garbo, Francesca & 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*, 15–60. Berlin: Language Science Press. DOI:10.5281/zenodo.3462778

Dobrushina, Nina. 1999. Glagol. In Aleksandr E. Kibrik (ed.), *Èlementy caxurskogo jazyka v tipologičeskom osveščenii*, 58–85. Moscow: MGU.

Donohue, Mark & Johanna Nichols. 2011. Does phoneme inventory size correlate with population size? *Linguistic Typology* 15(2). 161–170.

Evans, Nicholas. 1999. Why argument affixes in polysynthetic languages are not pronouns: Evidence from Bininj Gun-Wok. *Sprachtypologie und Universalienforschung* 52(3/4). 255–281.

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

Foley, William A. 1991. *The Yimas language of New Guinea.* Stanford: Stanford University Press.

Heath, Jeffrey. 1991. Pragmatic disguise in pronominal-affix paradigms. In Frans Plank (ed.), *Paradigms: The economy of inflection*, 75–89. Berlin: Mouton De Gruyter.

Heath, Jeffrey. 1998. Pragmatic skewing in 1<–>2 pronominal combinations in Native American languages. *International Journal of American Linguistics* 64(2). 83–104.

Hengeveld, Kees. 2012. Referential markers and agreement markers in Functional Discourse Grammar. *Language Sciences* 34(4). 468–479.

Janhunen, Juha. 2012. *Mongolian.* Amsterdam: John Benjamins.

Key, Harold H. 1967. *Morphology of Cayuvava.* The Hague: Mouton.

Kibrik, Aleksandr E. 2003. *Opyt morfologičeskoj rekonstrukcii kognitivnoj struktury (na materiale sfery ličnogo dejksisa v aljutorskom jazyke).* In: Aleksandr E. Kibrik: *Konstanty i peremennye jazyka*, 369–377. St. Petersburg: Aleteija.

Kibrik, Aleksandr E., Sandro V. Kodzasov & Irina A. Muravyova. 2004. *Language and folklore of the Alutor people* (Endangered Languages of the Pacific Rim A2-042). Osaka: ELPR.

Kibrik, Alexander E. 1985. Toward a typology of ergativity [selections]. In Johanna Nichols & Anthony C. Woodbury (eds.), *Grammar inside and outside the clause: Some approaches to theory from the field*, 268–323. Cambridge: Cambridge University Press.

Kibrik, Andrej A. 2011. *Reference in discourse.* Oxford: Oxford University Press.

Kulikov, Leonid. 2006. Case systems in a diachronic perspective: A typological sketch. In Leonid Kulikov, Andrej L. Malchukov & Peter de Swart (eds.), *Case, valency and transitivity*, 23–47. Amsterdam: John Benjamins.

Lacroix, René. 2009. *Description du dialecte laze d'Arhavi (caucasique du sud, Turquie): Grammaire et textes.* Université Lumière Lyon 2 dissertation.

Leer, Jeff. 1991. Evidence for a Northern Northwest Coast language area: Promiscuous number marking and periphrastic possessive constructions in Haida, Eyak, and Aleut. *International Journal of American Linguistics* 57(2). 158–193.

Liljegren, Henrik. 2019. Gender typology and gender (in)stability in Hindu Kush Indo-Aryan languages. In Francesca Di Garbo, Bruno Olsson & Bernhard Wälchli (eds.), *Grammatical gender and linguistic complexity: Volume I: General issues and specific studies*, 279–328. Berlin: Language Science Press. DOI:10.5281/zenodo.3462772

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

Mal'ceva, Alla A. 1998. *Morfologija glagola v aljutorskom jazyke*. Novosibirsk: Sibirskij xronograf.

Maslova, Elena. 2003a. *A grammar of Kolyma Yukaghir*. Berlin: Mouton de Gruyter.

Maslova, Elena. 2003b. *Tundra Yukaghir*. Munich: Lincom Europa.

Matasović, Ranko. 2014. Verbal and nominal agreement: Areal distribution and typological correlations. *Linguistic Typology* 18(2). 171–214.

Matasović, Ranko. 2007. The diachronic stability of gender revisited. Presented at Association for Linguistic Typology 8th Biennial Meeting, Paris.

Mattissen, Johanna. 2003. *Dependent-head synthesis in Nivkh: A contribution to a typology of polysynthesis*. Amsterdam: John Benjamins.

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

Nagayama, Yukari. 2003. *Ocherk grammatiki aljutorskogo jazyka* [Grammatical outline of Alutor]. (Endangered Languages of the Pacific Rim A2-038). Osaka: ELPR.

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

Nichols, Johanna. 2009. Linguistic complexity: A comprehensive definition and survey. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 110–125. Oxford: Oxford University Press.

Nichols, Johanna. 2011. *Ingush grammar*. Berkeley: University of California Press.

Nichols, Johanna. 2016. Complexity as non-canonicality: An affordable metric for morphology. SLE Annual Meeting, Leiden. (Submitted for Peter Arkadiev and Francesco Gardani, eds., *The complexities of morphology*.)

Nichols, Johanna. forthcoming. Canonical complexity. In Peter Arkadiev & Francesco Gardani (eds.), *The complexities of morphology*. Oxford: Oxford University Press.

Nichols, Johanna & Balthasar Bickel. 2013. Possessive classification. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/59.

Öztürk, Balkız & Markus A. Pöchtrager (eds.). 2011. *Pazar Laz*. Munich: Lincom Europa.

Robins, Robert H. 1958. *The Yurok language, grammar, texts, lexicon* (University of California Publications in Linguistics 15). Berkeley: University of California Press.

Schulze, Wolfgang. 1998. *Person, Klasse, Kongruenz. Fragmente einer Kategorialtypologie des einfachen Satzes in den ostkaukasischen Sprachen.* Vol. 1: *Die Grundlagen (1. Teil).* Munich: Lincom Europa.

Shosted, Ryan K. 2006. Correlating complexity: A typological approach. *Linguistic Typology* 10(1). 1–40.

Svantesson, Jan-Olof. 2003. Khalkha. In Juha Janhunen (ed.), *The Mongolic languages*, 154–176. London: Routledge.

Thurston, William R. 1982. *A comparative study in Anêm and Lusi.* Canberra: Australian National University.

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

Watkins, Laurel J. & Parker McKenzie. 1984. *A grammar of Kiowa.* Lincoln: University of Nebraska Press.

Wilson, William H. 1982. *Proto-Polynesian possessive marking.* Canberra: Australian National University.

Wood, Esther J. 2007. *The semantic typology of pluractionality.* University of California, Berkeley dissertation.

Yu, Alan C-L. 2003. Pluractionality in Chechen. *Natural Language Semantics* 11(3). 289–321.