

GENDER VARIATION ACROSS THE OROMO DIALECTS: A CORPUS-BASED STUDY*

Tekabe Legesse Feleke  & Terje Lohndal

Abstract. This study aims to (1) demonstrate the position of the Oromo gender system in Corbett's (1991) typology of gender; (2) illustrate major syntactic gender variation across the Oromo dialects; (3) identify factors that contributed to the gender variation, and (4) illustrate the morphosyntax of the Oromo gender system. The data obtained from the Oromo Speech Corpus shows a high degree of lexical and syntactic variation between the Western-Northern dialects on the one hand, and the Eastern-Southern dialects on the other hand. The Western and Northern dialects have shifted from the historically Cushitic phonology-based gender assignment pattern to the semantic-based assignment pattern. This shift has resulted in a widespread neutralization of feminine gender markers. The contact between Oromo and the neighboring non-Cushitic languages contributed to these changes. The study also argues that Kramer's (2015) morphosyntactic approach can be extended to the analysis of Oromo gender system.

1. Introduction

1.1. Background

Grammatical gender in Cushitic languages has received a fair amount of attention in the research literature due to its robust variation and complex interaction with number (see Mous 2008; Tsegaye 2017). However, the morphosyntax and dialectal variation when it comes to gender in Cushitic languages in general and Oromo in particular, have not received that much attention. As far as we are aware, Clamons (1993) is the only study that has been fully devoted to the dialectal variation in the Oromo gender system. Nevertheless, Clamons (1993) did not investigate the morphosyntax of Oromo gender assignment. Furthermore, his analysis of variation in Oromo gender assignment pattern targeted only a few dialects (mainly the Harar and Mecha dialects). More importantly, Clamons (1993) was conducted decades ago, but our theoretical understanding of grammatical gender and gender variation has changed substantially since then. Recently, Yadate (2019) also investigated the relationship between sexism and gender assignment in three languages: Oromo, Amharic and Gamo. The study viewed Oromo gender assignment from a social perspective, based only on data from the Wallaga (Mecha) dialect. Several recent studies of Oromo nominals

*We are grateful to the anonymous reviewers for their insightful comments. All remaining errors are ours. This work was supported by AcqVa Aurora Research Center at UiT, The Arctic University of Norway, grant number 2062165.

Studia Linguistica 2023, pp. 1–43. © 2023 The Authors. *Studia Linguistica* published by John Wiley & Sons Ltd on behalf of Editorial Board of *Studia Linguistica*.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Published by John Wiley & Sons Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK, and 350 Main Street, Malden, MA 02148, USA

2 Tekabe Legesse Feleke & Terje Lohndal

consistently omit a proper discussion of Oromo gender assignment and agreement (see Hawine 2009; Olani 2017; Wakweya 2017; Mewis 2001). Hence, how gender assignment is morpho-syntactically realized in Oromo and how the gender assignment and agreement systems vary across the Oromo dialects remain as an important object of further investigation. An important objective of the present contribution is to fill this lacuna.

The present study combines typological and theoretical approaches to show how gender is assigned in Oromo and to explore dialectal variations in Oromo gender assignment and agreement. First, it provides a detailed description of Oromo gender assignment patterns and illustrates the position of Oromo gender in Corbett's (1991) typology of gender. It argues that, in Oromo, gender assignment pattern varies across dialects; the Northern and Western Oromo dialects have predominantly a semantic-based gender assignment system while the Eastern and the Southern dialects have a phonology-based gender assignment pattern. It also reflects on the sources of this variation. In this regard, the study argues that a long history of contact between Oromo and Ethiopian Semitic languages is the main factor that underpins gender change in Oromo. Most importantly, the study argues that Kramer's (2015) morphosyntactic approach can be applied to analyze the morphosyntax of Oromo gender assignment. Specifically, in line with Kramer (2015), it argues that *n*, not the roots, is the locus of Oromo gender features. Since Kramer (2015) primarily focused on languages that have a semantic-based gender assignment such as Amharic, the present study extends Kramer's notion to languages and varieties that have a form-based gender assignment system. To this end, the present study provides a unified analysis of the morphosyntax of Oromo gender assignment and empirically illustrates the interface between the semantics and morphosyntax of gender. In particular, it shows that in languages that have a form-based gender assignment pattern, specific roots are licensed only by certain gender features, *ns*. Put together, the present study addresses four main objectives: (1) it determines the position of the Oromo gender system in Corbett's (1991) typology of gender; (2) it shows major semantic and syntactic gender variation across the Oromo dialects; (3) it identifies main factors that contribute to these variations, and finally (4) it analyzes the morphosyntax of the Oromo gender assignment. The study addresses these objectives based on data obtained from the Oromo Speech Corpus¹. The corpus consists of audio recordings and transcriptions of seven Oromo dialects: Arsi, Borana, Guji, Harar, Ilubabur, Kamise and Wallaga.

The main contents of the paper are organized as follows. Following this brief introduction, Section 1.2 presents some background information about the languages involved. Section 2 summarizes theoretical assumptions related to gender typology and to the morphosyntax of

¹ See Oromo Speech Corpus: <https://tekstlab.uio.no/glossa2/oromo>

gender. Section 3 provides a brief introduction to the Oromo gender assignment patterns. Section 4 shows gender assignment and agreement variations across the Oromo dialects and presents factors that contribute to the variations. Section 5 provides an analysis of the morphosyntax of the Oromo gender system. Finally, Section 6 presents conclusions and final remarks.

1.2. *The context of the study: Oromo and its varieties*

Oromo is a Lowland East Cushitic language. It is widely spoken in Ethiopia, but there are also Oromo speaking communities in Kenya and Somalia (see Clamons 1993; Clamons 1992; Olani 2017; Yimam 1986). The present study focuses only on the Oromo dialects that are spoken in Ethiopia. Oromo has many dialects, but the exact number of the dialects is unknown (cf. Blazek 2010; Heine, 2012; Kebede 2012; Mekonen, 2002; Negesse 2015). The present study's discussion regarding Oromo gender assignment and agreement variation relies on two recent Oromo dialect classification proposals: Negesse (2015) and Kebede (2012). Kebede (2012) classified Oromo dialects into West (*Shawa, Wallaga*), Central (*Arsi* (Rift Valley area), *Guji, Borana*), East (*Arsi* (Highland), *Harar*), *Kamise* and *Rayya*. Negesse (2015) proposed six sub-groupings of the Oromo dialect: Western dialect (*Wallaga, Jimma, Ilubabur*), Central dialect (*Shawa*), Northern dialect (*Rayya, Kamise*), South-eastern dialect (*Arsi, Bale*), Southern dialect (*Borana, Guji*) and Eastern dialect (*Harar*). The present study investigates seven Oromo dialects (*Arsi, Borana, Guji, Ilubabur, Harar, Kamise and Wallaga*) that are recognized in the two recent classification studies. *Shawa* and *Rayya* dialects are excluded since the Oromo Speech Corpus that the present study uses as the main source of data does not incorporate the two dialects.

Following Negesse (2015) and Kebede (2012), we divided the seven target dialects into four broad clusters: *Northern cluster* (Kamise), *Eastern cluster* (Harar), *Southern cluster* (Borana and Guji), and *Western cluster* (Arsi, Ilubabur and Wallaga) (see Figure 1). The Arsi dialect data analyzed in the present study were obtained from the *Zway Dugda* district, which is geographically close to the Western cluster. Hence, in Figure 1, the Arsi dialect is included within the *Western cluster*. Later, it will become apparent that the Arsi dialect of the *Zway Dugda* area indeed has a lot of similarity with dialects in the Western cluster. Except the Northern Kamise dialect, all dialects are spoken in the Oromia Regional State (see Figure 2), one of the states in Ethiopia. In Figure 2, Arsi can be taken as a point of reference to illustrate the areal distribution of the Oromo dialects. The Western dialect is located to west of Arsi, and the Eastern dialect is to the east of Arsi. The Southern dialect is spoken in the areas to the south of Arsi. The Northern dialect is spoken in the Amhara Regional State, outside the Oromia Region.

4 Tekabe Legesse Feleke & Terje Lohndal

The *white-blue* color in Figure 2 shows the administrative divisions of the Oromia Region. The figure shows that the Western dialects share a broad border with Amharic, Omotic and Nilo-Saharan languages. On the other hand, the Eastern and Southern dialects have a long history of contact with other Lowland East Cushitic languages, mainly with Afar and Somali. Later, it will be argued that the robust lexical and syntactic gender variation among the Oromo dialects is indeed the outcome of these contact situations. Regarding the sociolinguistic status of the Oromo dialects, currently there is no officially recognized *standard* Oromo dialect (Negesse 2015), regardless of several standardization efforts (see Demie 2010; Mekonen 2002; Tegegne 2015). Nevertheless, there is a consensus among scholars that the *Western* dialect is the one that has been dominant in schools and media outlets (see Tegegne 2015). The use of Oromo in schools and media outlets is a recent development (Beyessa 2014; Taye 2019). Since the early 1990s, Oromo has been taught in elementary schools in the Oromia Regional State, both as a subject and as a medium of instruction. Currently, Oromo is also taught as a subject in higher institutions in Ethiopia. It is also a working language of the Oromia Regional State. These changes might have significant ramifications regarding the gender variation across the Oromo dialects.

The Speech Corpus upon which the present study relies was developed by the University of Oslo, in collaboration with Addis Ababa University. The data extracted from the Oromo Speech Corpus was obtained from 88 adult native speakers who live in 18 different locations across the Oromia Region. As can be seen in Table 1, the Corpus consists of 266,

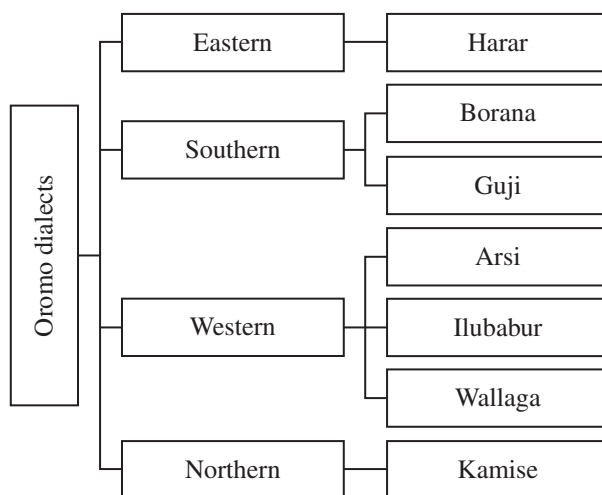


Figure 1. Classification of Oromo dialects, based on Negesse (2015) and Kebede (2012)

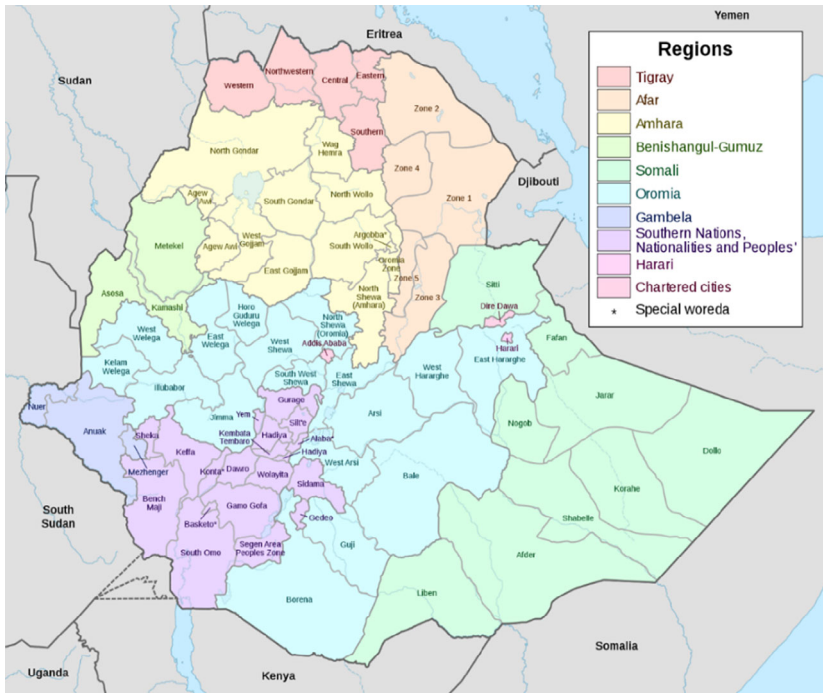


Figure 2. Oromo administrative map

500 tokens (see also <https://tekstlab.uio.no/glossa2/oromo>). Besides, it contains basic demographic information of the informants: sex and age. However, there is an asymmetry regarding data representation in the Corpus. For example, the data of the Arsi dialect was collected only from one locality (*Zway Dugda*) whereas the Guji dialect data was collected from six localities. There is also an asymmetry of tokens: the Borana dialect has the largest token number (54,416) while Arsi dialect has the smallest token number (9,020). We realize that these asymmetries may affect our conclusions and return to them in the discussion section.

2. Theoretical background

2.1. Gender variation and the typology of gender

Following Corbett (1991) and Hockett (1958), we define grammatical gender as classes of nouns reflected in the behavior of words that agree with nouns. Gender assignment is related to the way gender is allotted to different nouns, whereas gender agreement is the concord displayed on elements that agree with the noun (cf. Corbett 1991). Corbett (1991) discusses two types of gender assignment: *form-based* and *semantic-based*.

Table 1. General information about the Oromo speech Corpus

	Arsi (AR)	Borana (BR)	Guji (GJ)	Harar (HR)	Ilubabur (IL)	Kamise (KM)	Wallaga (WL)	Total
Locations	Zway- Dugda	<i>Bakke, Dabuluq, Mega, Nagelle & Yabello</i>	<i>Adoola, Bale-Horaa, Dama, Garba Anno, Goro Dola & Qarca</i>	<i>Harar & Gaaramullataa</i>	Mattu	Kamise	<i>Shambu & Nakamte</i>	18
Speakers	2	16	15	4	15	21	15	88
Tokens	9,020	54,416	51,878	11,969	37,289	81,001	21,008	266,500

In some languages, gender is assigned primarily based on the phonological or morphological form of the nouns while in some other languages, the semantics or natural gender plays a crucial role. In Section 4, we will see that Oromo dialects vary in terms of their gender assignment patterns: the Western and Northern Oromo dialects have a semantic-based gender assignment pattern. In other words, if the meaning of a noun is known, its gender can be predicted without reference to its form, either morphological or phonological (cf. Corbett, 1991:7–31). For example, in Western Oromo dialect, nouns denoting males such as *korma* ‘ox’ are always masculine while nouns denoting females such as *saʔaa* ‘cow’ are always feminine. In the case of same-root² inanimate and animate nouns, all the nouns receive masculine gender by default. The Eastern and Southern Oromo dialects have a phonology-based gender assignment system, i.e., gender assignment is established by reference to word-final vowels. In these dialects, every noun that ends with low central vowels /a/ or /aa/ has masculine gender while the rest of the nouns have feminine gender.

However, as noted by Corbett (1991), there is no system that is exclusively form-based, without any involvement of semantics. For instance, in the Eastern and Southern Oromo dialects where gender assignment is primarily phonology-based, kinship nouns such as *abbaa* ‘father.M’ and *haada* ‘mother.F’ are assigned gender according to semantic rules. This indicates that parts of the system are determined by the semantics, and that there can be overlap between the semantic-based and phonology-based gender assignment patterns. According to Corbett (1991), gender assignment can also be either overt or covert. Languages in which the gender of a noun is evident from its form are described as having *overt* gender; those where gender is not shown by the form of the nouns have *covert* gender. Related to this, in Section 4, we will show that the Western and Northern Oromo dialects have a covert gender assignment pattern since the gender of nouns appears only on verbs and adjectives via agreement. The Eastern and Southern Oromo dialects have overt gender marking, by virtue of gender being assigned according to the phonological form of the nouns.

Corbett (1991) shows that gender agreement targets adjectives, verbs, definite nouns, and various types of modifiers. We will illustrate in Section 3 and 4 that, in Oromo, gender agreement is expressed by inflectional suffixes that appear on verbs and adjectives. The only gender marking elements that appear in word-initial positions in Oromo are the historically Cushitic *t* and *k* as in *tiyyaa* ‘mine.F’ and *kiyyaa* ‘mine.M’. However, these markers are not prefixes; rather, they are part of the

² According to Kramer (2015), same-root nominals refer to nouns that are formed from the same root. For instance, in Oromo, *harree* ‘donkey.M’ and *harree* ‘donkey.F’ are same-root nominals.

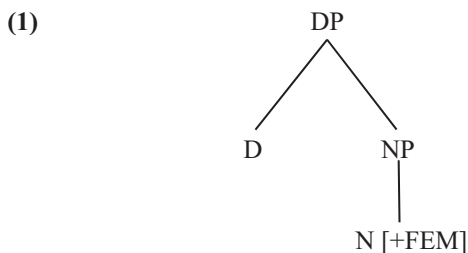
8 Tekabe Legesse Feleke & Terje Lohndal

roots. According to Corbett (1991), gender agreement can also be either *alliterative* or *non-alliterative*. In alliterative agreement, the formal gender marker on the noun is the same as the one used for agreement. We will see that the Oromo gender agreement pattern falls within the non-alliterative typology.

Corbett (1991) further stresses that gender assignment is an essential element particularly since native speakers must know the gender of the nouns to produce the correct agreement. Previous studies also indicate that gender assignment is a prerequisite for the acquisition of gender agreement (see Hopp 2013; Bates, Devescovi, Pizzamiglio, D'amico & Hernandez 1995; Marchal, Bianco, Dessus & Lemaire 2007). Relevant to the current discussion is the role gender assignment plays in understanding diachronic changes in gender agreement. In Section 4 we argue that neutralization in Oromo gender agreement is primarily triggered by irregularities in gender assignment (also see Duke 2010 for a similar argument). Furthermore, following Lohndal and Westergaard (2016), Dekeyser (1980), Enger (2004), Clamons (1992) and Igartua (2019), we illustrate the link between the diachronic change in gender assignment and the pattern of loss in gender agreement. We show that, in Oromo, gender loss is initiated primarily by language contact, and that the stability in a gender assignment system is determined by gender stability in the adjacent or nearby languages. Following Igartua (2019) and Trudgill (2013), we also discuss major factors that have probably resulted in gender loss in Oromo. Based on Igartua (2019) and Clamons (1992), we further contend that changes in gender assignment and agreement favor simplicity. We show that whenever gender loss occurs, it is the feminine gender that is primarily affected or neutralized.

2.2. *The Morphosyntax of gender*

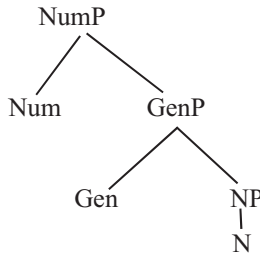
The simplest and perhaps oldest theory of grammatical gender is that it is a feature that is idiosyncratically listed in the lexicon on each noun, as illustrated in example (1). This approach is often considered *a fully listed approach* (see Kramer 2015).



However, as noted by Kramer (2015), this approach fails to capture the common generalization that male entities usually receive masculine gender and female entities receive feminine gender. In many languages, a noun can have either masculine or feminine gender, depending on the natural gender of the referent. For instance, in Oromo, *haakima* ‘doctor’ can be either masculine or feminine based on the biological gender of the referent. Hence, the fully listed approach fails to accommodate the fact that two semantically related homophonous nouns can have different gender. In other words, it neglects the fact that gender systems have a ‘semantic core’ (see Corbett, 1991); the idea that regular semantic-based gender assignment partly operates in all languages of the world. Kayne (2005) and Kramer (2015) also note that the fully listed approach assumes that gender features are located on roots. They argue that putting gender on roots make the root subject to crosslinguistic variation which contradicts the Chomsky-Borer hypothesis where variation is argued to be limited to features on functional heads.

More recent versions of the fully listed approach (see Roca 1989; Harris 1991; Ralli 2002; Riente 2003; Carstens 2010) make a distinction between natural gender (usually for animates) and grammatical gender. Accordingly, nouns are listed in the lexicon with either specified or unspecified grammatical gender; nouns with unspecified grammatical gender is assigned a grammatical gender via lexical rules that refer some semantic property of the nouns. Hence, inanimate nouns and animate nominals whose gender do not depend on the natural gender of the referent have their grammatical gender fully specified in the lexicon since their gender is not assigned on the bases of any semantic property. Hence, for example, Oromo different-root nominals such as *abbaa* ‘father’ and *haada* ‘mother’ have specified natural gender, but their grammatical gender is unspecified and must be assigned via lexical rules that connect lexical and grammatical gender. Kramer (2015) has challenged this line of reasoning. She argues that as far as the natural gender is the gender relevant for agreement, it is also present in the syntax, and it can trigger agreement directly. Given that the gender assignment pattern in Oromo also relies on natural gender (see Section 3 and 4), we assume that Kramer is on the right track. In Section 4, we will justify our position, based on Oromo data. Several other non-lexicalist models were also previously applied to the formal analysis of gender. For instance, Picallo (1991) proposed a morphosyntactic structure in which gender heads its own, *Gen(der)P*. According to this proposal, a gender feature can be merged below the Num head, which basically houses number inflection as illustrated in (2). *Gen* immediately dominates NP, and the *Gen* head is the source of gender inflection for all nouns.

(2)

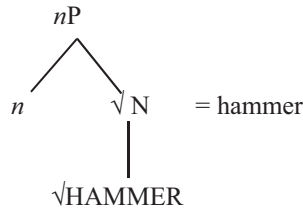


Arguments against *GenP* appears in Kramer (2015) and Ritter (1993). According to Kramer (2015), *GenP* is a projection that has no consistent semantics. In other words, there is always an uninterpretable arbitrary gender feature on inanimate nominals, and this lack of interpretability is not in line with Chomsky's (1995) dictum that projections in the syntax should not be based on uninterpretable features. Furthermore, according to Ritter (1993), the *GenP* analysis does not generalize since it cannot be applied to languages such as Hebrew in which gender affixes are derivational (a new nominal can be formed by adding, for example, a feminine suffix to a masculine inanimate noun). Ritter (1993) argues that in languages in which gender is inflectional, gender is housed in *Num* and not in *GenP* for two reasons: (i) for example plural nouns in Romanian switch gender, i.e., *Num* must have gender and number specifications, and (ii) in some languages, gender and number are spelled out together as the realization of a *Num* head, separately from a nominal head that is not inflected for either. In Oromo, there is no gender switching in plural nouns. Besides, as we will see in Section 3 and 4, derivational morphology does not result in gender change in Oromo. Hence, we will not pursue Ritter's (1993) line of argumentation.

As indicated in Section 1.1, Kramer's (2015) model which we adopt for the present study relies on the notion of Distributed Morphology and proposes a morphosyntactic approach which assumes that lexical categories are made up of a *category-neutral root* and a *category determining head* *n*, where *n* is the locus of the gender feature. The core argument in Kramer's (2015) proposal is that nouns such as *hammer* consist of the root $\sqrt{\text{HAMMER}}$ and the functional head *n* that nominalizes it, as illustrated in (3). Kramer (2015) argues that gender features should not be located on the root, but on *n*. We will see in Section 3 that in Oromo same-root nouns such as *haakim-ičča* 'doctor.M-PRT' and *haakim-itti* 'doctor.F-PRT' roots do not change their forms based on gender. Hence, guided by the Oromo data, we follow the claim that the root is not the locus of gender features. Kramer (2015) further argues that gender, as a phenomenon, is specific to a particular lexical

category (nouns) and putting category-specific information on roots undermines the idea that roots are category neutral. According to Kramer (2015), assigning gender to the root involves converting the root into a nominal (see also Acquaviva 2009). Matching the root to the right type of *n* requires certain licensing conditions so that a root is licit in the context of a given masculine or feminine *n*. For instance, *root-different*³ Oromo nouns such as *abbaa* ‘father’ and *haada* ‘mother’ which are assigned natural gender have licensing conditions. However, *same-root* nouns such as *harree* ‘donkey.F’ and *harree* ‘donkey.M’ do not have licensing conditions. In the latter case, the distinction between masculine *donkey* and feminine *donkey* is determined only based on discourse conditions. When the appropriate discourse context is not available, *same-root nouns* automatically receive the default gender, which is feminine in this case. Hence, the *n* analysis matches roots and gender through licensing conditions and accounts for both *same-root* and *different-root* nouns. Kramer (2015) proposes gender assignment rules which can be root-specific, i.e., different roots take different gender or different types of *n*. In her framework, morphophonological forms are inserted after syntax⁴.

(3)



Kramer (2015) shows that gender features come in two different types: interpretable for natural gender and uninterpretable for arbitrary gender. In Kramer’s analysis, there are four types of *ns* that are fully accountable for gender assignment to Amharic nominals.

³ According to Kramer (2015), *root-different* nominals refer to nouns that are formed from different roots.

⁴ We will see later that, in Oromo, gender affixes do not usually appear on nouns. The diminutive marker *-itti* as in for example, *man-itti* ‘a small house.F’ is the only gender-related suffix that appear on Oromo nominals. Based on Kramer (2015) and Oromo data, we argue that this suffix carries gender.

(4)

Types of *ns*

- a. *n* i [+FEM] Feminine natural gender
- b. *n* i [-FEM] male natural gender
- c. *n* no natural gender (natural gender irrelevant)
- d. *n* u [+FEM] Feminine arbitrary gender

According to Kramer (2015), the licensing conditions determine which root combines with which type of *n*. When same-root nominals such as *haakima* ‘doctor’ combine with (a), a female doctor is formed, and when they combine with (b) a male doctor is formed. A doctor whose gender is unknown is formed when they combine with (c). Female arbitrary gender on inanimates and feminine default gender are the result of an uninterpretable [+FEM] feature on *n* as illustrated by (d). Kramer (2015:51) proposes that the licensing conditions split into two sets: (i) licensing conditions that are related to semantics and (ii) licensing conditions that are arbitrary (unrelated to any property of the root, *n*). A different-root nominal is licensed under one of the interpretable *ns*. These licensing conditions are semantic since the *n* that a root combines with has an impact on the semantic interpretation. In other words, it causes *nP* to be interpreted as male or female. We will see that same-root Oromo nominals and feminine default animate nouns can combine with both interpretable *ns* (4a and 4b). Thus, there is no need to specify any semantic licensing conditions for these nouns. The gender features of inanimate nouns are generally considered to be irrelevant for semantic interpretation. Arbitrary licensing conditions that are not root specific apply to these nouns. In other words, they match a particular root with a particular *n*. Unlike the semantic licensing conditions, this matching is

Table 2. Oromo consonant inventory

Consonants		Labials	Alveolars	Palatals	Velars	Glottals
Stop	VL	p	t	č	k	ʔ
	VD	b	d	gč	g	
	Ejective	pʼ	tʼ	cʼ	kʼ	
	Implosive		ɗ			
Continuants	VL	f	s	š		h
	VD		z			
Nasal		m	n	ñ		
Liquid			l,r			
Glides		w	y			

Table 3. Oromo vowel inventory

	Front	Central	Back
High	i		u
Central		e	o
Low		a	

idiosyncratic, *i.e.*, a chance association between a root and an *n*. Our main contribution in this regard is, contrary to Kramer's argument, the matching between roots and the feature *n* is not always idiosyncratic. Rather, in languages that have form-based gender assignment, there is a clearly predictable association between the form of the roots and the feature of *n*. We further illustrate that in Oromo the diminutive determines the agreeing gender. That is, diminutivized Oromo nominals have multiple *ns*, and the highest *n* determines the agreeing gender of the nominals.

3. The Oromo gender system: An Overview

Oromo is a Lowland East Cushitic language. According to Baye (1983:9), the phonemic inventory of Oromo has 25 consonants (see Table 2). Each consonant has a long or geminate counterpart as for example in *hatuu* 'to cheat' and *hattuu* 'thief'. The only exceptions in this regard are /ʔ/ and /h/ as these do not have geminate counterparts.

Table 3 shows that Oromo has five vowel phonemes. Vowel length is contrastive in Oromo; for example, *gala* 'return' and *gaala* 'camel'. As a tradition, vowel length is denoted by double vowels. We adopted this approach in the present study.

Like many other Afroasiatic languages, Oromo has two gender classes: *masculine* and *feminine* (see Clamons 1992, 1993; Mous 2012). In most cases, the gender of kinship terms is assigned based on the natural gender of the referent (see examples in (5)). These terms are described as *different-root* nouns in Kramer (2015)'s analyses.

(5)

a. abbaa, 'father.M' vs. haadā 'mother.F'

b. ilma, 'son.M' vs. intala 'girl.F'

c. obboleessa 'brother.M' vs. obboleetti 'sister.F'

d. adaadaa 'aunt.F' vs. adeera 'uncle.M'

e. akaakoo 'grandfather.M' vs. akkoo 'grandmother'

f. ḡaarsa 'husband.F' vs. ḡaartii 'wife.F'

There are also kinship terms whose gender cannot be determined based on the natural gender. These kinship terms have masculine default gender. Some examples are provided in (6).

14 Tekabe Legesse Feleke & Terje Lohndal

(6)

- | | |
|---|---|
| a. soddaa 'son -in-law or sister-in-law' | c. k'ut'usuu 'youngest brother or sister' |
| b. angafa 'elder brother or elder sister' | d. fira 'male/female close family member' |

In Oromo, nouns referring to natural creations usually receive feminine gender as examples in (7) illustrate.

(7)

- | | |
|---------------------|--------------------|
| a. daččii 'earth:F' | c. baatii 'moon.F' |
| b. aduu 'sun.F' | d. urġii 'star.F' |

(8)

- | | |
|----------------------|--------------------------|
| a. k'alamee 'ape.F' | f. sinii 'cup.F' |
| b. harree 'donkey.F' | g. harc'c'umme 'stick.F' |
| c. gaangee 'mule.F' | h. buk'k'ee 'gourd.F' |
| d. reʔee 'goat.F' | i. k'aršii 'money.F' |
| e. aroo 'spider.F' | j. t'aafii 'teff.F' |

In Oromo, only a handful of nouns end in consonants. These nouns have masculine gender as the examples in (9) illustrate.

(9)

- | | |
|-------------------------|--------------------------------|
| a. midaan 'food item.M' | f. finc'aan 'urine.M' |
| b. torbaan 'week.M' | g. fuññaan 'nose.M' |
| c. loon 'cattle.M' | h. abaar 'draught.M' |
| d. bišaan 'water.M' | i. hiraar 'frustration.M' |
| e. halkan 'night.M' | j. mi'taat'is 'sweet potato.M' |

In Oromo, if the biological gender of a noun is unknown, masculine gender is assigned by default.

(10)

- | | |
|---|--|
| a. namuu dūfuu barbaad-ø-a
everyone to come want-M-IPFV
'Everyone wants to come.' | b. namni hundi raf-ø-e
everybody sleep-M-PFV
'Everybody slept' |
|---|--|

The gender assignment pattern of the remaining nouns varies across dialects. In the Eastern and Southern dialects, gender is assigned according to the phonological form of the nouns, *i.e.*, nouns that end

in the low central vowel /a/ and /aa/ receive masculine gender while nouns that have other vowel endings receive feminine gender. In the Western and Northern dialects gender is assigned based on semantic parameters. Nouns that do not have natural gender receive masculine default gender. This issue will be discussed in Section 4.

In Oromo, gender is expressed via agreement. A morphological marker of gender appears on nouns only when the nouns are diminutivized. Hence adding *-itti* to an inanimate masculine noun results in a diminutive interpretation (*small, cute, elegant*) of the nominal. In (11), *mukatti*, for example, refers to a small-sized wood with feminine gender.⁵

(11)⁵

- a. muka ‘wood.M’ vs. mukatti ‘diminutivized wood.F’
- b. mana ‘house.M’ vs. manatti ‘diminutivized house.F’
- c. fallaana ‘spoon.M’ vs. fallaanatti ‘diminutivized spoon.F’
- d. laga ‘river.M’ vs. lagatti. ‘diminutivized river.F’

The diminutive nominals trigger feminine agreement in Oromo (see examples in (12)), *i.e.*, the diminutive determines the agreeing gender. Following Kramer (2015), in Section 5, we will analyze the Oromo diminutive as an *n* head that takes an *nP* complement.

(12)

- | | |
|---|---|
| <ul style="list-style-type: none"> a. man-ittii-n tun t’innoo dà house-DIM.F-NOM this.F small COP ‘This house is small.’ | <ul style="list-style-type: none"> b. muk-ittii-n c’ab-t’-e wood-DIM.F-NOM break-3.F.SG-PFV ‘The wood has broken’ |
|---|---|

Oromo gender agreement targets verbs and adjectives. The examples in (13) illustrate gender agreement between nouns and verbs. The agreement is indicated by the gender suffixes that appear on the verbs. There are some differences among the Oromo dialects regarding the agreement between plural nouns and verbs. These differences will be discussed in Section 4.

(13)

- | | |
|--|--|
| <ul style="list-style-type: none"> a. harree-n duf-t-e donkey-NOM come-3.F.SG-PFV ‘The donkey came.’ | <ul style="list-style-type: none"> c. waraaabeess-i ni-duf-- Ø-a hyena-NOM FOC-come-3.M.SG-IPFV ‘The hyena will come/comes.’ |
| <ul style="list-style-type: none"> b. gurbaa-n duf-Ø- e boy-NOM come-3.M.SG-PFV ‘The boy came.’ | <ul style="list-style-type: none"> d. intalt-i ni- duf-t-i girl-NOM FOC-come-3.F.SG-IPFV ‘The girl will come/comes.’ |

⁵ The feminine form *-itti* expresses both specificity and diminution. For example, *man-itti* can mean either a particular house or a small-sized.

16 Tekabe Legesse Feleke & Terje Lohndal

The examples in (13) also indicate that, in Oromo noun-verb gender agreement, only feminine gender is overtly marked. The masculine gender is marked by a default empty morpheme. In Oromo, adjectives also agree with nouns in gender. There is gender agreement, for example, between nouns and possessive adjectives (14a, b), nouns and relative pronouns (14c), nouns and interrogative pronouns (14d), and nouns and demonstrative pronouns (14e).

(14)

- | | |
|--|---|
| <p>a. mana kiyya
house.M my.M
'My house'</p> <p>b. gaangee-n tun tiyya
mule.F-NOM this.F mine.F
'This mule is mine.'</p> | <p>c. holaa kan ati bitte
sheep.M which.M you bought
'The sheep which you bought'</p> <p>d. intala tam?
gil.F which.F
'Which girl?'</p> <p>e. gurbaa kana
boy.M this.M
'This boy'</p> |
|--|---|

According to Owens (1985), there are five types of descriptive adjectives in Oromo: (1) those that are not marked for gender, which include adjectives such as *adii* 'white', *gaari* 'generous', *dihoo* 'near', *ñaliila* 'pure', *booru* 'unpurified', *k'ulk'ulluu* 'clean', *guutuu* 'full' and *miskiina* 'pity'; (2) those that are marked by *-aa* and *-oo* for masculine and feminine gender, respectively. Adjectives within this category include *guddaa* 'big.M' / *guddoo* 'big.F', *ñik'k'aa* 'small.M' / *ñik'k'oo* 'small.F', *dip'p'aa* 'narrow.M' / *dip'p'oo* 'narrow.F'; (3) those that are distinguished by *-aa* or *-tuu* suffixes to mark masculine and feminine gender respectively as in for example, *hamaa* 'dangerous.M' / *hamtuu* 'dangerous.F', *ğabaa* 'clever.M' / *ğabduu* 'clever.F', *bareedaa* 'hudson.M' / *bareedduu* 'beautiful.F', *fokkotaa* 'ugly.M' / *fokkottuu* 'ugly.F'. (4) those that are distinguished by *-a* and *-tii* suffixes to mark masculine and feminine gender respectively as in, for example, *dureessa* 'rich.M' / *dureetti* 'rich.F', *hiyyeessa* 'poor.M' / *hiyyeetti* 'poor.F'. The last group of adjectives are distinguished by *-içça* and *-itti* to mark masculine and feminine gender respectively, for example *gurraaçça* 'balck.M' / *gurratti* 'balck.F', *dullacçaa* 'old.M' / *dullatti* 'old.F'. The examples in (15) illustrate gender agreement between nouns and some selected descriptive adjectives. There is variation among the

Table 4. Oromo dialects

	Western	Southern		Eastern	Western	Northern	Western
Dialect name	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
Abbreviation	AR	BR	GJ	HR	IL	KM	WL

Oromo dialects concerning agreement between nouns and adjectives. This issue will be the focus of the next section.

(15)

- | | |
|---|---|
| <p>a. intal-tt-i c'im-tu-n k'o?at-t-e
girl-F-NOM clever-F-FOC study.3.F.SG-PFV
'The clever girl studied.'</p> <p>b. gurbaa-n deeraa-n k'o?at-Ø-e
boy-NOM tall.M-FOC study-3.M-PFV
'The tall boy studied.'</p> | <p>c. inn-i barataa c'imaā dā
he-NOM student clever.M COP
'He is a clever student.'</p> <p>d. iṣee-n saffis-ttu dā
she-NOM fast-3.F.SG COP
'She is fast.'</p> |
|---|---|

Examples (4)–(15) show that both semantic and form-based gender assignment systems exist in Oromo. Grammatical gender is assigned based on biological gender in the dialects, including in dialects in which gender assignment is primarily form-based. The Eastern and Southern Oromo dialects use an overt gender assignment system whereas the Northern and Western Oromo dialects use a covert gender assignment system. In the next section, we capitalize on this issue.

4. Gender variation across the Oromo dialects

Now we elaborate on gender assignment and agreement variations and on the factors contributing to the variation by providing examples from the Oromo Speech Corpus. Section 4.1 presents the variation in Oromo gender assignment. Section 4.2 deals with variation in Oromo gender agreement. Finally, Section 4.3 briefly describes factors that have contributed to the gender variation. The seven Oromo dialects are summarized in Table 4.

4.1. Variation in gender assignment.

As indicated in Section 3, the Easter and Southern Oromo dialects are different from the Western and Northern dialects in terms of gender assignment. The examples in (16) illustrate gender assignment in the Arsi (AR) dialect.

(16)

- | | | |
|--|---|---|
| <p>a. barsiisaa 'teacher.M'</p> <p>b. kitaaba 'book.M'</p> <p>c. baruumsa 'education.M'</p> <p>d. gaara 'mountain.M'</p> | <p>e. waraana 'war.M'</p> <p>f. harka 'hand.M'</p> <p>g. yeroo'time.M'</p> <p>h. horii 'animal.M'</p> | <p>i. oduu'news.M'</p> <p>j. naannoo 'area.M'</p> <p>k. hattuu 'thief.M'</p> <p>l. bakkee 'place.M'</p> |
|--|---|---|

18 Tekabe Legesse Feleke & Terje Lohndal

Examples in (16) show that, nouns in the Arsi dialect receive masculine gender regardless of the phonological differences in word-final position. Examples in (16 g-l) contain nouns that end in non-low central vowels, /a/ and /aa/. However, these nouns take masculine gender. It should be noted that the Arsi data presented here was obtained from the *Zway Dugda* area, a place which is geographically very close to the Western dialect. We have limited data available for the Arsi dialect.

Unlike the Arsi dialect, gender assignment in the Borana (BR) dialect is purely phonology-based. The examples in (17) illustrate that nouns that end with the low central vowel /a/ and /aa/ take masculine gender while nouns that end with other types of vowels receive feminine gender.

(17)

a. biyyee ‘soil.F’	e. dubbi ‘speech:F’	i. gaafa ‘hornM’
b. iḡoollee ‘child.F’	f. intala ‘girl:F’	j. guyyaa ‘day.M’
c. baallii ‘feather.F’	g. goočča ‘action.M’	k. holaa ‘sheep.M’
d. dabballee ‘cadre.F’	h. ilaalčča ‘perspective.M’	l. aadaa ‘culture.M’

Examples in (17a-e) show that nouns that end in non-low central vowels receive feminine gender while examples in (17 g-l) show that nouns that end in the vowel /a/ and /aa/ receive masculine gender. Example (17f) indicates that natural gender overrides the grammatical gender in the Borana dialect. As ‘girl’ is biologically female, it receives feminine gender even though it ends in the low central vowel /a/.

Like the Borana dialect, the Guji (GJ) Oromo dialect has a phonology-based gender assignment pattern as illustrated in (18). The examples in (18a-f) present nouns that end with the non-low central vowel. These nouns occur with feminine gender. Nouns that end in the vowel /a/ and /aa/, (18 g-l), receive masculine gender.

(18)

a. gurmuu ‘shoulder.F’	c. rakkoo ‘challenge.F’	e. gurmuu ‘unity.F’
b. dubbii ‘speech.F’	d. ilmee ‘child.F’	f. gabbii ‘heifer.F’
g. dāla ‘generation.M’	i. gaafa ‘horn.M’	k. deega ‘tradition.M’
h. mučaa ‘boy.M’	j. handuura ‘umbilicus.M’	l. torbaan ‘week.M’

However, many of the nouns in the Ilubabur (IL) dialect receive masculine gender. Examples in (19a-f) present nouns that end in the vowels /a/ and /aa/. These nouns receive masculine gender. The remaining nouns also receive masculine gender though they have non-/a/ endings.

(19)

- | | | |
|-------------------------|------------------------|---------------------|
| a. k'eerransa 'tiger.M' | e. morma 'neck.M' | i. horii 'animal.M' |
| b. dádaa 'butter.M' | f. sirba 'song.M' | j. zayitii 'oil.M' |
| c. geerarsa 'song.M' | g. yeroo 'time.M' | k. killee 'egg.M' |
| d. gogaa 'leather.M' | h. manguddoo 'elder.M' | l. k'awwee 'gun.M' |

Like Borana and Guji dialects, the Harar (HR) dialect has a phonology-based gender assignment system. In the Harar dialect, nouns that end in the low central vowels /a/ and /aa/ receive masculine gender (20 g-l) while those that end in other vowels receive feminine gender (20a-f).

(20)

- | | | |
|-----------------------|---------------------------|-------------------------|
| a. šittoo 'perfume.F' | e. naannoo 'area.F' | i. gurbaa 'boy.M' |
| b. buk'k'ee 'gourd.F' | f. badii 'mistake.F' | j. ŋaata 'food.M' |
| c. šeekkoo 'sheik.F' | g. c'absaa 'marriage.M' | k. hoḡḡaa 'local tea.M' |
| d. yeroo 'time.F' | h. garaabaa 'khat leaf.M' | l. nama 'human.M' |

Like the Arsi and Ilubabur dialects, the Kamise (KM) dialect has masculine default gender; most of the nouns take masculine gender irrespective of their vowel endings. Examples in (21a-f) illustrate the gender of nouns that do not have /a/ endings whereas examples in (21 g-l) present the gender of nouns that have /a/ endings.

(21)

- | | | |
|--------------------------|-------------------------|---------------------|
| a. yaroo 'time.M' | e. dallansuu 'curse.M' | i. mak'aa 'name.M' |
| b. oduu 'news.M' | f. k'allabi 'ration.M' | j. dimma 'issue.M' |
| c. rakabotii 'chassis.M' | g. adaa 'culture.M' | k. bakka 'place.M' |
| d. hoḡii 'job.M' | h. sokkaa 'departure.M' | l. malkaa 'river.M' |

The Wallaga dialect has also masculine default gender assignment pattern. Examples in (22a-f) illustrate nouns that have /a/ and /aa/ endings whereas examples in (22 g-l) present nouns that have non-/a/ endings. Both sets of nouns receive masculine gender.

(22)

- | | | |
|---------------------------|-----------------------|-----------------------|
| a. buna 'coffee.M' | e. yaada 'idea.M' | i. gahee 'share.M' |
| b. adaa 'culture.M' | f. ŋaata 'food.M' | j. iḡollee 'child.M' |
| c. dáloota 'generation.M' | g. yeroo 'time.M' | k. bakkee 'place.M' |
| d. biyya 'country.M' | h. rakkoo 'problem.M' | l. šamarree 'girls.M' |

The above examples show that in the Eastern (Harar) and the Southern (Borana and Guji) Oromo dialects, nouns that end with low central vowels, /a/ and /aa/, receive masculine gender while the remaining nouns receive feminine gender. However, there are a few exceptions to this. The examples in (23) present some of the nouns that receive feminine gender even though they have /aa/ endings (also see Temam, 2020).

(23)

- | | |
|-------------------------|------------------------------|
| a. bobaa ‘armpit.F’ | f. guuftaa ‘hair cover.F’ |
| b. bortaa ‘force.F’ | g. haxuuraa ‘present gift.F’ |
| c. dállaa ‘back bald.F’ | h. madaa ‘wound.F’ |
| d. dūgaa ‘fact.F’ | i. moraa ‘camp.F’ |
| e. fooddaa ‘window.F’ | j. tooftaa ‘tactic.F’ |

Examples (16)–(23) show that in the Eastern (Harar) and the Southern (Borana and Guji) dialects, gender is assigned based on the phonological form of the nouns whereas in the Western (Arsi, Ilubabur and Wallaga) and in the Northern (Kamise) dialects, most nouns receive masculine default gender. Besides the default gender, in Oromo, nouns may also be assigned gender based on various pragmatic conditions, *gender switch* due to the sensitivity towards a change in the discourse conditions. The most common conditions are the condition of the size of an object and the condition of the speakers’ knowledge of the biological gender of the referent. For example, *saree* ‘dog’ may have feminine default gender, but after realizing the biological sex of the dog, native speakers may switch to masculine gender *saree* ‘dog.M’. Likewise, *mana* ‘house’ may have masculine default gender in all the dialects, but a small-sized ‘house.F’ *manitti* can receive feminine gender. Given the examples provided in (16)–(23), and the kinship terms discussed in Section 3 (see example (4 & 5)), we assume that the following gender assignment rules apply to the Western and Northern Oromo dialects.

(24)

- a. If a noun refers to a male referent, masculine gender is assigned
- b. If a noun refers to a female referent, feminine gender is assigned
- c. If a nominal referent has no or unknown natural gender, masculine gender is by default assigned.

The rules in (24) predict that in these dialects gender is assigned based on semantic factors. If the semantic factors are absent, masculine gender is assigned by default. Note that these rules exclude exceptional inanimate nominals that have default feminine gender (see example (7) & (8)) and gender assignment which is triggered by the above-mentioned two discourse conditions. The present study does not fully address context-driven gender switch in Oromo. It only touches up on it to illustrate the complexity in Oromo gender assignment system. Figure 3 shows a comprehensive gender assignment schema for the Western and Northern Oromo dialects. The *t* and *k* illustrated in the Figure are historically Cushitic gender marking elements. The proposed schema incorporates both the default gender (including nouns with default feminine gender) and the gender of discourse referents.

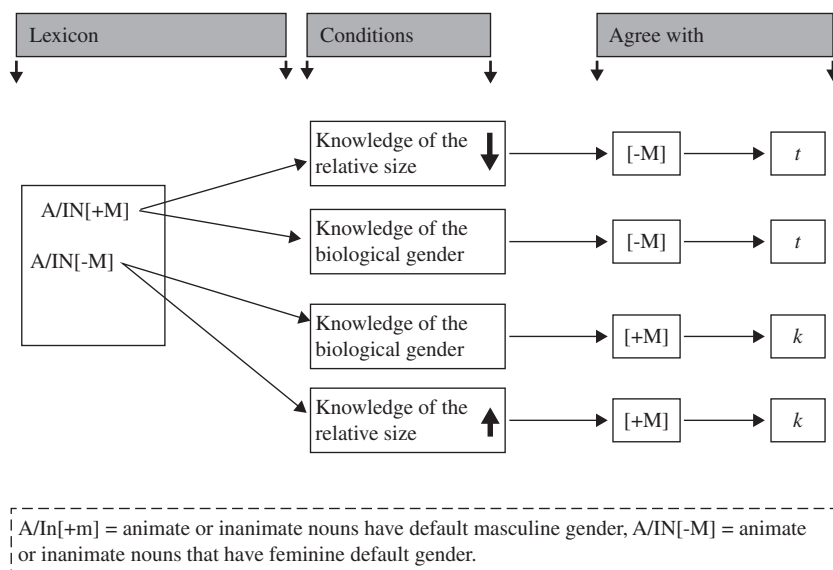


Figure 3. Gender assignment schema for the Western and the Northern Oromo dialects

Figure 3 shows that every noun⁶ is assigned either feminine [-M] or masculine [+M] gender based on the rules listed in (24); for example, *horse.M*, *mule.F*, *house.M* and *gourd.F*. Then inanimate nouns which lack biological gender, such as *buk'k'ee* 'gourd' and *mana* 'house' are further constrained by the discourse condition of size. Hence, *house* may receive feminine gender if its size is extremely small, for example *manittii* 'a small house', although its default gender is masculine. The same is true for inanimate nouns that have feminine default gender such as *gourd*: native speakers switch to the masculine gender if the *gourd* has a relatively big size. Likewise, animate nouns that have default masculine gender, for example *farda* 'horse', may receive feminine gender under *the condition of the speakers' knowledge of the biological gender of the referent*, for example *farditti* 'a female horse'. In other words, once the speakers realize that the animate nominal is biologically female, they automatically assign the feminine gender. Likewise, a nominal that has a feminine default gender such as *gaangee* 'mule', can receive masculine gender once its biological gender is identified as male, *gangičča* 'a particular male mule'. Kinship terms that have natural gender such as *mother.F*, *father.M* and *sister.F* remain unaffected; that is to say, they are constrained neither

⁶ We are using the traditional framework just for descriptive purposes. In Section 5 we will illustrate that, in Kramer's morphosyntactic approach, gender assignment happens in the syntax.

22 Tekabe Legesse Feleke & Terje Lohndal

by the discourse condition of size nor by the condition of the speakers' knowledge of the biological gender.

Examples in (16)–(23) further illustrate that, in the Eastern (Harar) and Southern (*Borana* and *Guji*) Oromo dialects, gender is assigned based on the phonological form of the nouns. It was also illustrated in Section 3 that gender assignment rules for the kinship terms and for the nouns that have feminine default gender are invariable across the Oromo dialects. Given all this, we assume that the following gender assignment rules apply to the Eastern and Southern dialects.

(25)

- If a noun refers to a male animate referent, masculine gender is assigned.
- If a noun refers to a female animate referent, feminine gender is assigned.
- If a nominal referent ends with the low central vowel /a/ or /aa/, masculine gender is assigned.
- Otherwise, feminine gender is assigned (*elsewhere condition*).

Figure 4 illustrates gender assignment schema of the Eastern and Southern Oromo dialects. Once gender is assigned based on the rules listed in (25), the two discourse conditions are applied to the nouns to satisfy conversation contexts. In the Eastern dialect, for example, *saree* 'dog' has feminine default gender by virtue of ending in a non-low central vowel, /e/, but after realizing that a particular dog is biologically male,

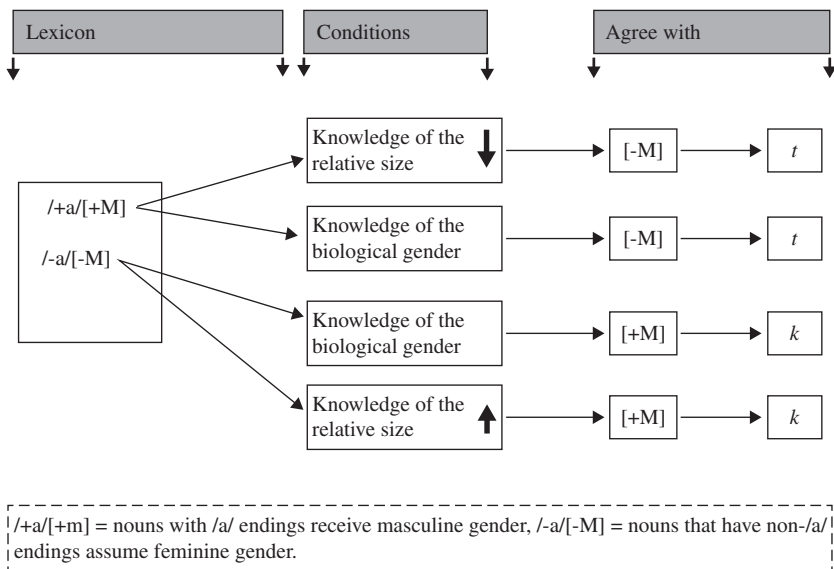


Figure 4. Gender assignment schema for the Eastern and the Southern Oromo dialects

native speakers switch to masculine gender: *saree* ‘dog.M’. The reverse is true for nouns that have masculine default gender such as *farda* ‘horse’. Likewise, the base gender may shift due to the size of the inanimate referent. For example, *mana* ‘house’ has a masculine default gender since it ends with vowel /a/, but a small-sized house, *manitti*, receives feminine gender.

4.2. Variation in gender agreement

In Section 3, it was indicated that Oromo gender agreement targets two agreement areas: the agreement between nouns and adjectives, and the agreement between nouns and verbs. This section discusses the extent to which these two agreement types vary across the Oromo dialects. Table 5

Table 5. summary of gender variation across the Oromo dialects

No	Areas of variation	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
1	Phonology-based system	*	✓	✓	✓	*	*	*
2	Noun-possessive pronoun agreement	*	✓	✓	✓	*	*	*
3	Noun-relative pronoun agreement	*	✓	✓	✓	*	*	*
4	Noun-interrogative pronoun agreement	✓*	✓	✓	✓	*	*✓	*
5	Noun-proximal demonstrative agreement	✓*	✓	✓	✓	*	*✓	*
6	Collective nouns-SG verb agreement	✓	✓	✓	✓	✓	✓	✓
7	PL Noun-PL demonstrative	*✓	*	*	*✓	*✓	*✓	*✓
8	Gender-neutral adjectives	*	*	✓	*	✓	✓	✓
9	Plural noun-singular verb agreement	*✓	✓	✓	✓	*✓	*✓	*✓

Note: ✓ = exists, *✓ = sometimes exists; * = does not exist

presents the summary of gender variations across the Oromo dialects that we will discuss. The table shows that *Arsi*, *Kamise*, *Ilubabur* and *Wallaga* dialects are different from the remaining three dialects both in terms of gender assignment and gender agreement.

Oromo shows gender agreement between nouns and varieties of adjectives, mainly pronouns, and descriptive adjectives. In general, many of the masculine pronominals in Oromo contain the historically Cushitic masculine gender marker *k* while the feminine pronominals contain the Cushitic feminine gender marker *t*. Table 6 presents possessive pronouns in the seven Oromo dialects. As can be seen from the table, the feminine possessive pronouns contain the feminine marker *t* (e.g. *tiyya* ‘my.F’) while the masculine possessive pronouns contain the masculine marker *k* (e.g., *kiyya* ‘my.M’). Table 6 further indicates that there is cross-dialectal variation regarding the type of possessive pronouns used across the dialects. The Eastern and Southern dialects have maintained both the feminine and masculine forms while only the masculine form has survived in the Northern and Western dialects. In the table, the feminine and masculine possessive pronouns are separated by a slash. In the Western and Northern dialects, where only the masculine possessive pronouns survived, the possessive pronouns are presented without the

Table 6. Oromo possessive pronouns (feminine/masculine)

	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
1 st	kiyya (my)	tiyya/ kiyya (my)	tiyya/ kiyya (my)	tiyya/ kiyya (my)	kiyya (my)	kiyya (my)	kiyya (my)
	keñña (our)	teñña/ keñña (our)	teñña/ keñña (our)	teñña/ keñña (our)	keñña (our)	keñña (our)	keñña (our)
2 nd	kee (your)	tee/kee (your)	tee/kee (your)	tee/kee (your)	kee (your)	kee (your)	kee (your)
	kessan (your)	tessan/ kessan (your)	tessan/ kessan (your)	tessan/ kessan (your)	kessan (your)	kessan (your)	kessan (your)
3 rd	(ka) isaa (his)	(ta/ka) isaa (his)	(ta/ka) isaa (his)	(ta/ka) isaa (his)	(ka) isaa (his)	ka isaa (his)	(ka) isaa (his)
	(ka) iŕsee (her)	(ta/ka) iŕsee (her)	(ta/ka) iŕsee (her)	(ta/ka) iŕsee (her)	(ka) iŕsee (her)	ka iŕsee (her)	(ka) iŕsee (her)
	(ka) isaani (their)	(ta/ka) isaani (their)	(ta/ka) isaani (their)	(ta/ka) isaani (their)	(ka) isaani (their)	ka aani (their)	(ka) saani (their)

slash. Later, we will see that a few feminine pronominals are available in each of the Northern and Western Oromo dialects.

In the Eastern and Southern dialects, feminine nouns agree in gender with the feminine possessive pronouns while masculine nouns agree with the masculine possessive pronouns, for example *mana.M kiyya.M* ‘my house’ and *intala.F tiyya.F* ‘my girl’ (see also Gemeda 2006). However, in the Northern and Western dialects, every possessed noun agrees just with the masculine possessive pronouns. In other words, agreement between feminine nouns and feminine possessive pronouns has been neutralized in these dialects. There are a few remnants of gender agreement between feminine possessed nouns and feminine possessive pronouns in the Northern and Western Oromo dialects as illustrated in (26). Nevertheless, this agreement occurs with nouns and pronouns whose antecedents have natural gender.

(26)

- a. muc’ayyo-n tiyya haataatu ġa-tt-e (AR) b. iṣee-n tiyya adda (KM)
 girl-NOM my.F be said-3.FSG-PFV she-NOM my.F different
 ‘Lit. My girl agreed’ ‘Mine is different’.

Regarding relative pronouns, like the possessive pronouns, the Eastern and Southern dialects have preserved the *t* vs. *k* distinction. However, in the Western and Northern dialects, only the masculine form, *k*, has survived. A few feminine relative pronouns are found in Arsi, *Guji*, *Wallaga* and *Kamise* dialects, but they are insignificant in number. Table 7 presents relative pronouns in the seven Oromo dialects.

In the Eastern and Southern dialects, masculine nouns agree in gender with the masculine relative pronouns, and feminine nouns agree with the feminine relative pronouns. However, in the Northern and Western dialects, only the masculine form, *kan*, is used with both feminine and masculine nouns. The examples in (27) illustrate gender agreement between nouns and relative pronouns.

Table 7. Oromo relative pronouns

	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
SG.M	kan	kan	kan	kan	kan	kan	kan
SG.F	kan	tan	tan	tan	kan	kan	kan

(27)

- a. adaa tan tiyyeefanna barabaaddu (BR)
culture.F that.F attention need.F
'Culture that needs attention.'
- b. ilmee tan akaakootaa laatani (GJ)
child.F that.F grandfathers.BEN give-IMF
'A child that is given to the grandfathers.'
- c. dalagaa tan yaroo deer-tu fudat-tu (HR)
work.F that.F time long-F take-3.F.SG
'Work that takes a long time.'
- d. c'absii kan ġir-Ø-u (IL)
traditional marriage.F that.M exist-3.M.SG-IPF
'Traditional marriage that exists.'
- e. yoo kan dabartu taate (KM)
If that.M pass.F happen.F
'Lit: If she passes.'
- f. dubartii-n kan kaleessa duf-t-e (WL)
girl-NOM that.F yesterday come-3.F.SG-PFV
'A girl that came yesterday.'

The same holds true for the interrogative pronouns. Both masculine and feminine interrogative forms are available in the Eastern and Southern dialects, but only the masculine form is frequently used in the Northern and Western dialects. Table 8 shows the type of interrogative pronouns that are used in the seven Oromo dialects. In Arsi (Western) and Kamise (Northern) dialects, both the feminine and masculine forms are detected, but the feminine interrogative pronoun is used only with nouns that receive natural feminine gender.

Likewise, both the masculine and feminine proximal demonstrative pronouns survived in the Eastern and Southern dialects while only the masculine one is available in the Western and Northern dialect. Table 9 shows a wide range of syncretism among the Oromo demonstrative pronouns.

In general, Oromo dialects do not make a distinction between masculine and feminine distal demonstrative pronouns. Hence, the same form *san* 'that' is used both with masculine and feminine nouns; for example, *intala san* 'that girl' or *gurbaa san* 'that boy'. Furthermore, the same form of distal demonstrative pronouns is used with both singular and plural nouns: *barataa san* 'that boy' and *barattoota san* 'those boys'.

Table 8. Oromo interrogative pronouns

	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
M	kam	kam	kam	kam	kam	kam	kam
F	kam/tam	tam	tam	tam	kam	kam/tam	kam

Table 9. Oromo demonstrative pronouns

Demonstratives	Arsi	Borana	Guji	Harar	Ilubabur	Kamise	Wallaga
Prox.SG (F/M)	tun/ kun	tun/ kun	tun/ kun	tun/ kun	kun	tun/ kun	kun
Distal.SG	san	san	san	san	san	san	san

Regarding the proximal demonstratives in the Eastern and Southern dialects, distinct forms for masculine and feminine demonstratives exist as indicated in (28).

(28)

- | | |
|--|--|
| <p>a. <i>kun barataa.M dā</i>
 this.M student.M COP
 ‘This is a student’</p> | <p>e. <i>tun barattuu dā</i>
 this.F student.F. COP
 ‘This is a student’</p> |
|--|--|

In the Northern and Western dialects, however, only the masculine form survived. Hence, these dialects use the masculine, *kun*, both with masculine and feminine nouns. Some dialects have different forms of demonstrative pronouns (see Gemeda 2006). For instance, the proximal demonstrative pronoun can be realized either as *kun* or *kana*. Likewise, the distal demonstrative pronoun has three surface realizations: *tun*, *sun* and *san*. Additionally, nominative forms of the demonstrative pronouns always add nominative suffixes, either *-in* or *-i*. Hence, *kun* can surface either as *kunin* or *kuni*. Examples in (29) illustrate the agreement between nouns and demonstrative pronouns. The feminine demonstrative *tun* was also found in Arsi (Western) and Kamise (Northern) dialects, but it occurs very infrequently and occurs only with nouns that have female biological gender.

(29)

- | | |
|--|--|
| <p>a. <i>nami-čč-i kun ačči bah- Ø-a (AR)</i>
 man-PRT-NOM this.M there vacate-3.M.PIFV
 ‘This man vacates from there.’</p> <p>b. <i>kadaa-n tun ni taa-t-i (BR)</i>
 prayer-NOM this.F FOC happen-3.F.SG-PFV
 ‘Lit: This prayer will succeed.’</p> <p>c. <i>intal-t-i tun heerum-t-e (GJ)</i>
 girl-3.F.SG-NOM this.F marry-3.F.SG-PFV
 ‘This girl married.’</p> | <p>d. <i>intal-ti tun-i ġir-t-i (HR)</i>
 girl-PRT this.F-NOM survive-3.F.SG-PFV
 ‘Lit: This girl survived.’</p> <p>e. <i>k’awwee-n kun hiika-m- Ø-e (IL)</i>
 gun-NOM this.M disassemble-PASS-3.F.SG-PFV
 ‘This gun was disassembled.’</p> <p>f. <i>elaa-n tun ni heerum-t-i (KM)</i>
 girl-NOM this.F FOC marry-3.F.SG-IPFV
 ‘This girl will marry.’</p> <p>g. <i>intal-t-i kun-i waa beek-t-i (WL)</i>
 girl-3.F.SG this.M-NOM thing know-3.F.SG-IPFV
 ‘This girl knows something.’</p> |
|--|--|

The plural demonstrative pronoun, *kanneen*, is found in five of the seven dialects: *Arsi*, *Harar*, *Wallo*, *Ilubabur* and *Wallaga*. However, its frequency is insignificant (14 instances in the Kamise dialect out 23 total instances). Its feminine counterpart, *tanneen*, was not attested in any of the dialects. Given that the Corpus used in the present study contains only about 260, 000 tokens, it is difficult to conclude that the plural demonstrative pronouns have entirely vanished, but there is a clear indication that they are being replaced by singular demonstrative pronouns. The examples in (30) illustrate agreement between nouns and plural demonstrative pronouns.

(30)

- | | |
|--|---|
| <p>a. iḡoollee fi'tik'k'oo kanneen fuud-an-i (AR)
children.PL young.PL those marry-PL.PFV
'Those married young children'</p> <p>b. biyy-oota gudad-an kanneen akka ameerikaa (IL)
country-PL develop-PL those like America
'Lit: Those developed countries such as USA.'</p> | <p>c. kanneen k'oteebult-oota ta'ani (KM)
Those farmer-PL be.PL
'Those that are farmers'</p> <p>d. barat-oota kanneen darbb-an (WL)
student-PL that.M pass-PL-PFV
'Those passed students'</p> |
|--|---|

The above qualitative data show that the historically Cushitic feminine marker *t* has been neutralized in the Northern and Western Oromo dialects. A similar argument was previously made by Clamons (1992). Table 10 illustrates the percentage of the gender markers *t* and *k* that are found in the pronouns of each dialect. The table shows that the feminine marker *t* has almost completely eroded in the Western and Northern dialects. It seems that, in the *Arsi* dialect, it has partially survived. The partial survival of *t* in the *Arsi* dialect may be because the *Arsi* data analyzed in the present study was collected from the border area between the Eastern and Western dialects (see Section 2). In general, the results presented in Table 10 indicate that the feminine marker *t* has been overtaken by the masculine marker *k* in four of the seven dialects investigated in the present study.

Unlike the agreement between nouns and pronouns, there is only some cross-dialectal variation regarding the agreement between nouns and descriptive adjectives. As shown in Section 3, adjectives always agree in gender with singular nouns. The examples in (31) illustrate this.

(31)

- | | |
|---|---|
| <p>a. yeroo gabaabaa (AR)
time.M short.M
'A short time'</p> <p>b. ḡaar-ti-i laaf-tu (BR)
woman-F-NOM weak-F
'A weak woman'</p> <p>c. ḡireeñña dansaa (GJ)
life.M good.M
'A good life'</p> <p>d. ñaata bareedaa (HR)
food.M nice.M
'A nice food'</p> | <p>e. foon bareedaa (IL)
meat.M nice.M
'A nice meat'</p> <p>f. dubbii cima (KM)
issue.M serious.M
'A serious issue'</p> <p>g. yaada gabaabaa (WL)
idea.M short.M
'A short idea'</p> |
|---|---|

In many Oromo dialects, especially in the Southern and Eastern dialects, plural nouns agree in gender and number with plural adjectives. The plural adjectives are formed via *reduplication* as illustrated in (32) (also see Mohammed & Andrzej 1990). For instance, *bareedaa* 'handsome' is singular masculine while *babbareedaa* 'handsome' is plural masculine.

Table 10. The frequency of *t* and *k* observed across the Oromo dialects

	Total Frequency	<i>t</i> - Frequency	<i>k</i> - Frequency	% of <i>t</i>	% of <i>k</i>	Remark
Arsi	424	17	407	4.0	96.0	Neutralized
Borana	2751	759	1992	27.6	72.4	Preserved
Guji	1678	199	1479	11.9	88.1	Preserved
Harar	584	149	435	25.5	74.5	Preserved
Ilubabur	1321	4	1317	0.3	99.7	Neutralized
Kamise	4199	52	4147	1.2	98.8	Neutralized
Wallaga	759	3	756	0.4	99.6	Neutralized

(32)

- | | |
|---|--|
| <p>a. iḡooll-ee t'it'ik'k'oo (BR)
child-PL small.F.PL
'Small children.'</p> | <p>b. daa'imm-an t'it'ik'k'oo (GJ)
baby-PL young.F.PL
'Young baby'</p> |
| | <p>c. mann-een babareedaa (HR)
hous-PL beautiful.M.PL
'Beautiful houses'</p> |

In Oromo, adjectives that end with *-oo* usually indicate feminine gender (see Section 3). In Guji, Kamise, Wallaga and Ilubabur dialects, however, adjectives that have *-oo* endings are gender-neutral. We assume that in these dialects, the feminine marker *-oo*, is in the process of losing its feminine feature. Such gender-neutral adjectives are usually used with both singular and plural nouns, see examples in (33). Note that in the Harar dialect, for instance, *namoota beekamoo* 'well-known people' is ungrammatical. In this instance, only adjectives marked for gender are acceptable, *namoota beekamaa.M.*

(33)

- | | |
|--|--|
| <p>a. namoota čičimoo (GJ)
individual-PL clever.PL
'Clever people'</p> | <p>c. rakkina cimoo (KM)
challenge serious
'Serious challenge'</p> |
| <p>b. gos-oota babbeekamoo (IL)
tribe-PL well-known.PL
'Well-known tribes'</p> | <p>d. nam-oota beekkamoo (WL)
individual-PL well-known
'Well-known people'</p> |

In Oromo, it seems that number agreement between plural nouns and adjectives is optional since there are also singular nouns that are used with plural adjectives as illustrated in (34).

30 Tekabe Legesse Feleke & Terje Lohndal

(34)

- | | |
|--|---|
| <p>a. dākaa diddiimaa (AR)
stone red.M.PL
'Red stones'</p> <p>b. nama ḡaḡḡabaa (BR)
man strong.M.PL
'Strong men'</p> <p>c. minseensa t'it'ik'k'aa (GJ)
member small.M.PL
'Small members'</p> | <p>d. ḡollee babareedduu (HR)
child.F beautiful.F.PL
'Beautiful children'</p> <p>e. hoḡii gurguddaa (IL)
job great.M.PL
'Great jobs'</p> <p>f. nama gurguddaa (WL)
man tall.M.PL
'Tall men'</p> |
|--|---|

In Oromo, when two or more adjectives modify a noun, all the adjectives agree in gender with the nouns they modify as illustrated in (35). This is true for all dialects.

(35)

- | | |
|--|--|
| <p>a. gurbaa diimaa bareedaa
boy light-skinned.M handsome.M
'Light-skinned handsome boy'</p> <p>b. intala diimttuu bareedduu
girl light-skinned.F beautiful.F
'Light-skinned beautiful girl'</p> | <p>c. barataa dēeraa c'imaā saffisaa
student.M tall.M clever.M fast.M
'Fast clever tall student'</p> <p>d. barattuu dēerttuu cimttuu saffistuu
student.F tall.F clever.F fast.F
'Fast clever tall student'</p> |
|--|--|

In most cases, plural nouns agree in number with plural verbs, for instance *isaan barat-an* 'they learned' where *-an* is a third person plural marker. In this case, there is no overt gender agreement between nouns and verbs. However, noun-verb agreement in Oromo is somehow flexible, and in many of the dialects plural nouns can agree in gender with singular verbs as illustrated in (36).

(36)

- | | |
|---|---|
| <p>a. nam-oota rakkoo beek-t-u (AR)
people-PL challenge realize-3.F.SG-IPV
'People that realize a challenge.'</p> <p>b. dubrat-oota ollaa kee ḡir-t-u (BR)
girl-PL neighbor your.M exist-3.F.SG-IPV
'Lit: Girls that are your neighbors.'</p> <p>c. sabaa-fi sablamm-oota ḡir-t-u hunda (GJ)
nation-COJ nationality-PL exist-3.F-IPV all
'All existing nation and nationalities'</p> <p>d. dubar-toota naannoo kee jir-t-u (HR)
girl-PL around your.M find-F.IPV
'Lit: Girls around you.'</p> | <p>e. nam-oota walga-h-e keesstii (IL)
people-PL gather-3.M.SG-PFV in
'In the gathered people'</p> <p>f. šamarr-a-n karaa-irra teess-i KM)
girl-PL-NOM road-LOC sit.3.F.SG-IPFV
'Girls that sit on the road'</p> <p>g. nam-oota buna kana kal-ø-u (WL)
people-PL coffee this.M prepare-M.SG-IPV
'People that prepare this coffee'</p> |
|---|---|

Whether the plural nouns agree with feminine or masculine gender in the Southern and Eastern dialects depends on the phonological form of the number marking suffixes as example (37) illustrates. If the number suffixes contain /a/ or /aa/, the noun receives masculine gender, but if the noun ends with other vowels or consonants, it receives feminine gender.

We assume that in these dialects, the number suffixes serve as declension classes that specify the gender class of the plural nouns.

(37)

- | | |
|---|--|
| <p>a. ḡabbi-lee-n teeñña duf-t-e.
heifer-PL-NOM our.F come-F-PFV
'Our heifers came.'</p> <p>b. horr-oota-ni issaa deem-Ø-e.
donkey-PL-NOM his go-2.SG-F-PFV
'His donkeys came.'</p> | <p>c. galm-awwan-Ø keeñña iḡaara-m-an-Ø
hall-PL-NOM Our.M build-PASS-PL-PFV
'Our Halls were built.'</p> <p>d. hool-oota-ni isaami duf-Ø-e
sheep-PL-NO. their come-2.M.SG-PFV
'Sheep came.'</p> |
|---|--|

Figure 5 shows the dialect map of the seven Oromo dialects. The blue color shows the dialect area of the Eastern (Harar) and Southern (Borana and Guji) dialects which have a phonology-based gender assignment pattern. The orange color indicates dialect areas of the Western (Arsi, Ilubabur and Wallaga), and the Northern (Kamise) dialects which have a semantic-based gender assignment pattern.

4.3. Sources of gender variation

In Section 4.1 and 4.2, it was shown that gender assignment and agreement vary across the Oromo dialects. It was also indicated that the neutralization of the feminine marker *t*, the presence of gender-neutral adjectives, plural noun-singular verb agreement, and syncretism in varieties of pronouns are the major areas of variation. This section explains some of the factors underlying this variation. We argue that language contact, i.e., the contact between Oromo and neighboring

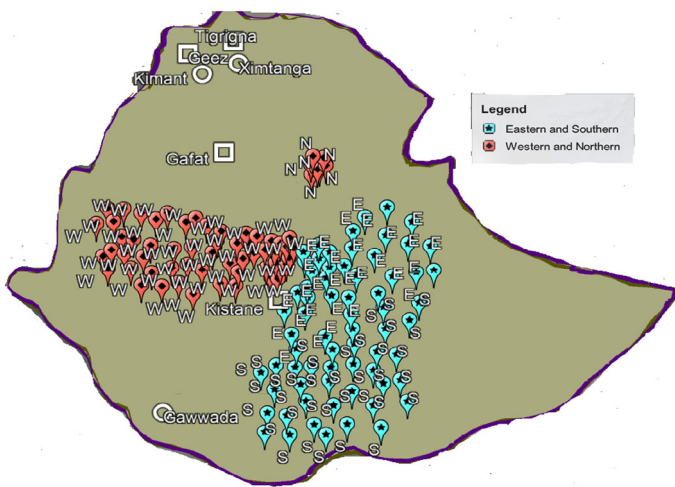


Figure 5. Oromo dialect areas

languages is the main factor that underpins gender variation in Oromo. This contact significantly influenced the Western and Northern dialects because of a long history of contact between these dialects and the Ethiopian Semitic languages (also see Clamons 1992). The *Northern* and *Western Oromo* dialects have centuries of history of contact with Ethiosemitic languages such as Amharic, Argobba and Tigrigna (see Appleyard 2015; Blench 2013; Crass & Meyer 2011; Feleke 2021; Feleke, 2023; Leslau 1945), and with Omotic languages (see Clamons 1992; Mous 2012). Of all these, the contact between Amharic, and the Western and Northern Oromo dialects is extremely relevant given that Amharic and Oromo also share an extended boundary (see Figure 2) and that they are spoken by millions of speakers. On the other hand, the Eastern and Southern dialects have had less contact with the Semitic languages, but they have been in contact with other Lowland East Cushitic languages, mainly with Afar and Somali (see Woldemariam & Lanza 2014; Zaborski 1991). Hence, it is likely that this contact has contributed to preserving the originally Cushitic gender assignment and agreement systems of the Eastern and Southern Oromo dialects.

Besides the areal contact, there are several instances of history of migration and expansion that make the contact between the Northern and Western Oromo dialects and Amharic crucial. In this regard, a notable historical event is the expansion of the Semitic speakers from the north to the south during the reign of King Amdetsion, in the early 14th century. At the time, the Semitic speakers expanded their territory from the north towards the south and occupied areas that had been under the control of non-Semitic speakers, mainly under the control of the Oromos (see Marcus 2002). This was followed by a reverse south-to-north Oromo movement in the early and late 16th century. Following this movement, the Oromo speakers, mainly the Mecha and Tulama tribes (speakers of the Western dialect), expanded from the south towards the north and west of Ethiopia and settled in areas that were previously occupied by non-Oromo speakers (see Hassen 2015; Clapham 1988). There was also a third massive expansion of Amharic speakers, from north towards the south, west and east in the late 19th century, during the reign of *Minilik II*. This movement resulted in a nation-wide spread of Amharic and in the introduction of Amharic as the National Language of Ethiopia. Undoubtedly, these movements and expansions might have resulted in the shift of the gender system of the Northern and Western Oromo dialects. They might have resulted in the absorption of the Cushitic gender system by semantic-based Semitic gender assignment pattern. The history of the Eastern and Southern dialects is quite different. The early and late 16th century Oromo movement also expands towards the east, and the Oromos occupied areas that had been under the control of various non-Oromo tribes. This event, combined with centuries of contact between the two Oromo dialects and the two widely spoken

Lowland East Cushitic languages (Somali and Afar), might have preserved the Cushitic gender system of the Eastern and Western Oromo dialects (see Banti 1988; Clamons 1992; Kefale 2010).

There is also linguistic evidence that substantiates the above claims. First, Ethiopian Semitic languages are known for having a semantic-based gender assignment system (see Gasser 2010; Kramer 2015; Leslau 1995; Leslau 2000). For example, in Amharic, gender is assigned mainly based on the natural gender of the referent (see Kramer 2015). Contrary to the sister Cushitic languages, the gender assignment patterns of the Western and Northern Oromo dialects match the Amharic semantic-based gender system. This shift must be due to the influence of Ethiopian Semitic given that other Lowland Cushitic languages such as Afar and Somali have phonology-based gender assignment (see Hayward & Corbett 1988; Oomen 1981; Kraska 1992; Saeed 1999). Moreover, the non-Lowland Cushitic languages that have been in contact with Oromo do not have a phonology-based gender assignment system. Hence, we assume that the phonology-based gender assignment is originally the system of proto-Lowland East Cushitic, and that the Western and Northern Oromo dialects shifted towards the Semitic gender system (see Clamons 1992 for similar arguments). The emergence of the gender-neutral adjective in the Western and Northern Oromo dialects could also be the consequence of Oromo-Ethiosemitic contact. Many Ethiopian Semitic languages, including Amharic, do not manifest gender agreement between nouns and predicative adjectives (see Kramer 2009; Kramer 2015; Workneh 2011). However, in most Cushitic languages, adjectives are always marked for gender (see Mous 2008; Tosco 2000).

The Eastern and Western Oromo dialects maintained the phonology-based gender assignment system although Amharic has been taught as a second language in areas where the two dialects are spoken. This implies that only intense and durable contact leads to gender loss. As illustrated in Section 4.2, only feminine gender is affected. According to Hetzron (1980) and Dekeyser (1980), when gender change is necessary, it is the feminine gender that is likely to be neutralized. A similar conclusion was reached by Lohndal and Westergaard (2016), Dekeyser (1980), Enger (2004), Clamons (1992), Igartua (2019) and Corbett (1991). Alpher (1987) further argues that in a patriarchal society, such as the Ethiopian society, the tendency of generalization towards the masculine gender is very high. Other studies have also previously reported a generalization towards the masculine gender in communities where multiple varieties are in close contact with each other (see Busterud, Lohndal, Rodina & Westergaard 2019; Bull and Swan 2002; Enger 2004; Lohndal & Westergaard 2016; Rodina & Westergaard 2015).

In general, we presume that gender change in Oromo involves two steps. First, the gender assignment pattern of the Western and Northern dialects shifted from the phonology-based to a semantic-based gender

assignment pattern. Then, the shift in gender assignment led to the neutralization of gender agreement. It is crucial to note that the shift in gender assignment is from a phonology-based two-gender system to a system which has default masculine gender. This means that gender assignment contributes to the diachronic changes in gender agreement (see Corbett 1991 & Matasović 2004 for similar argument). Clamons (1992) further argues that a change in gender assignment in general favors oversimplification. Hence, the neutralization of gender features in the Western and Northern dialects is the outcome of a shift from the phonology-based complex Cushitic gender system towards the simpler semantic-based Semitic gender system.

5. The morphosyntax of Oromo gender: A theoretical analysis

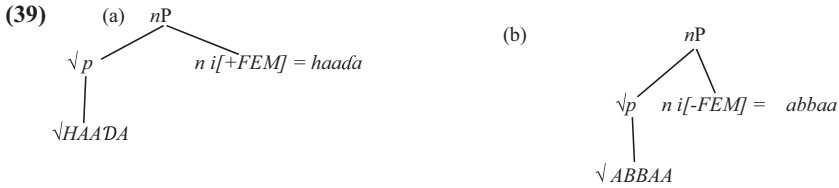
We will now turn to an analysis of the morphosyntax of gender assignment. We begin with the Northern and Western Oromo dialects that have the same gender system as Amharic, the language that Kramer (2015) builds her proposal upon. Like Amharic, the Western and Northern Oromo dialects have default masculine gender except the kinship terms and a few exceptional nouns that have feminine default gender. In section 3, it was shown that *different-root* nouns such as *haada* ‘mother’ and *adadaa* ‘aunt’ have natural feminine gender. According to Kramer (2015), these nouns have interpretable feminine *n* (38a), given that they are naturally female. A feature is assumed to be interpretable if its presence or absence changes the interpretation of a linguistic structure, i.e., if it is legible at LF (Kramer 2015:58). On the other hand, nouns such as *abbaa* ‘father’ and *oboleessa* ‘brother’ have natural masculine gender. These nouns have interpretable masculine *n* (38b). Moreover, *same-root* nouns such as *mana* ‘house’ and *farda* ‘horse’ are licensed under plain *n* since they do not have natural gender (37c) or they have default masculine gender. Finally, *same-root* nouns that have default feminine gender such as *aduu* ‘sun’ and *gaangee* ‘mule’ have ‘unlicensed’ feminine *n* (38d). The types of gender features in (38) provide a complete list of possible types of *ns* of the Western and Northern Oromo dialects.

(38) Types of *ns* (Northern and Western Oromo dialects)

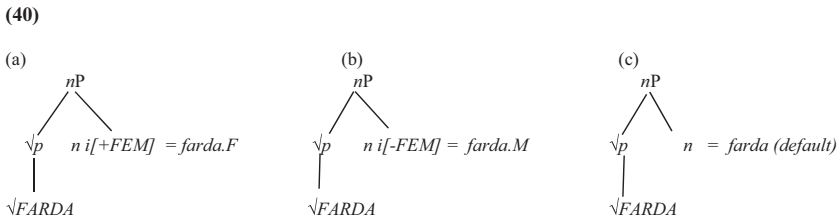
- a. *n* i [+FEM] Female natural gender
- b. *n* i [-FEM] Male natural gender
- c. *n* No natural gender or plain *n* (just default masculine gender)
- d. *n* u [+FEM] Female arbitrary gender

Summarizing, an *n* with an interpretable feature, i [+FEM], is interpreted as female natural gender, a *n* with an interpretable feature i [-FEM] is

interpreted as male natural gender, and a *n* with no gender feature is interpreted as having no natural gender. In (39a) and (39b), *haada* ‘mother’ and *abbaa* ‘father’ are licensed under *i* [+FEM] and *i* [-FEM], respectively.



As can be seen from the examples in (40), the *same-root* animate noun, for example *farada* ‘horse’, is licensed under one of the three conditions in (38). A horse that has female natural gender is licensed under [+FEM] as illustrated by (40a); a horse that has male natural gender is licensed under [-FEM] as illustrated by (40b), and a horse whose natural gender is unknown is licensed under plain *n* as illustrated by (40c). A horse whose natural gender is unknown receives masculine gender by default. Note that the selection of a contextually appropriate gender is determined by discourse or context.



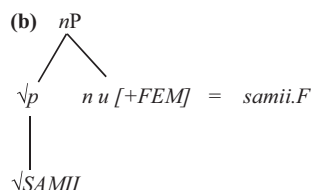
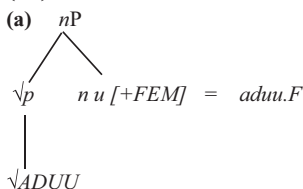
The analyses presented in (40) can also be applied to inanimate *same-root* nouns such as *mana* ‘house’ and *gaara* ‘mountain’ that have default masculine gender. When the roots MANA and GAARA combine with (38c), they form *house* and *mountain* whose genders are default masculine (see (41 a & b)).



Now, the remaining nouns are inanimate and animate nouns that have default feminine gender. Inanimate feminine nouns such as *aduu* ‘sun.F’

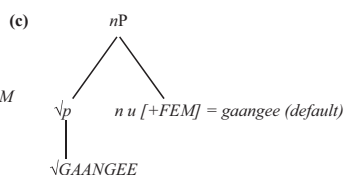
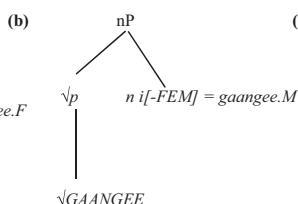
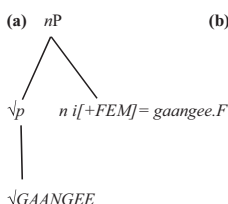
and *samii* ‘sky.F’ are licensed under the uninterpretable feature n, u [+FEM] (38d). Hence, these nouns surface with the same kind of agreement as naturally female animate nouns because they have a [+FEM] feature except that the feature is uninterpretable (see (42 a & b)). Note that these nouns are naturally feminine, and they do not receive masculine gender under any circumstance.

(42)



Regarding the animate nouns with feminine default gender such as *gaangee* ‘mule’, they are licensed under the interpretable *ns* ((38a) and (38b)) when their natural gender is known (see (43a-b)). However, when the natural gender is unknown or irrelevant, they are licensed under the uninterpretable u i[+FEM] (see (38c)).

(43)



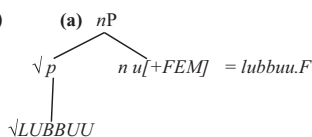
Now, regarding the Eastern and Southern Oromo dialects, in Section 3 it was illustrated that *root-different* nouns such as *haada* ‘mother’ and *abbaa* ‘father’ have natural gender. These genders are licensed precisely by the same conditions illustrated in (39a) and (39b). Hence, we do not illustrate that again here. Likewise, inanimate nominals that have default feminine gender such as *aduu* ‘sun’ and *samii* ‘sky’ are licensed by the uninterpretable condition illustrated in (42). The remaining nouns are now those whose gender are assigned based on the phonological form of the nouns. As illustrated in Section 4, the difference between the Eastern and Southern phonology-based, and the Western and Northern semantic-based gender system resides in these nouns. If any of these nouns ends with a low central vowel /a/ or /aa/, it receives masculine gender. Otherwise, it receives feminine gender. Note that according to Kramer (2015), different roots are licensed under different gender

features, *ns*. To accommodate this, we add the following two *ns* to the ones provided in (37).

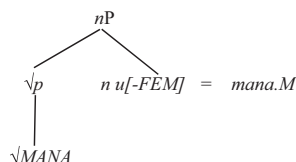
(44)

- a. *n u* [-FEM] for /a/ and /aa/ endings
 b. *n u* [+FEM] for the rest (elsewhere condition)

(45)



(b)



Example (45a) indicates that nouns that have non-/a/ vowel endings such as *lubbuu* ‘life’ are licensed under *n u* [+FEM]; hence, they receive feminine gender. On the other hand, example (45b) shows that nouns that end with the low central vowel /a/ and /aa/ such as *mana* ‘house’ are licensed under *n u* [-FEM]. Therefore, they receive masculine gender. These examples show that the association between the root and the feature on *n* is not necessarily arbitrary; only specific roots i.e., those that end with the low vowels /a/ and /aa/ are licensed under [-FEM]. In the same manner, only roots that have non-/a/ endings are licensed under [+FEM].

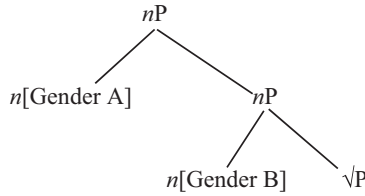
It was indicated in Section 3 that, in Oromo, the diminutive determines the agreeing gender. Following Kramer (2015), we argue that the diminutive is formed via *n* selecting for a nominal complement. We further assume that diminutivized nominals have multiple *ns*. In nominals that have multiple *ns*, the highest *n* determines the agreeing gender of the nominal, i.e., the gender feature on the diminutive morpheme determines the gender of the nominals. There are two competing explanations for this, one based on the property of hierarchical structure, and the other based on cyclicity (see Kramer 2015 for an explanation based on hierarchical structure). Cyclicity makes the following three assumptions:

(46)

- a. *n* is the phrase head, i.e., head of the cyclic domain
 b. A phase head triggers the spell-out of any cyclic domain they c-command
 c. Phase Impenetrability Condition

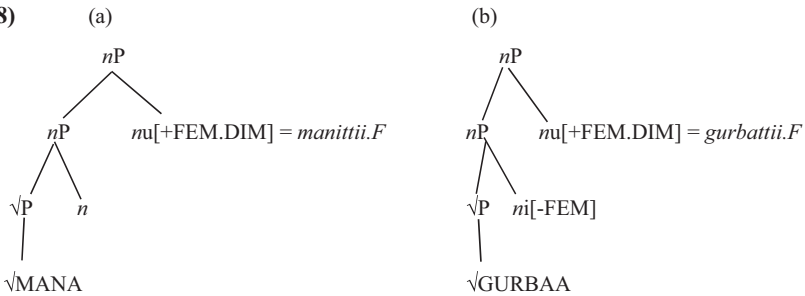
In (47), an *nP* with Gender A will trigger the spell-out of its complement, the lower *nP*, since the lower *nP* is a cyclic domain. The lower *nP* will then be inaccessible to later syntactic operations. This means that only the higher *nP* is accessible for agreement since the lower *nP* will have been sent to the interface before any agreement targets have been merged.

(47)



Therefore, in Oromo, masculine inanimate nouns such *mana* ‘house’ receive a diminutive interpretation illustrated in (48a) when it is used with feminine agreeing forms. Animate nouns such *gurbaa* ‘boy’ receive a diminutive interpretation illustrated in (48b) when it is used with feminine agreeing forms.

(48)



In (48a) and (48b), the diminutive is analyzed as an *n* with an uninterpretable feminine gender feature, *n* *u*[+FEM]. In (48a), since *mana* ‘house’ is a masculine inanimate noun, it is licensed under a plain *n*. On the other hand, in (48b), *gurabaa* ‘boy’ is a root-different animate noun and therefore, it is licensed under an interpretable *n*[-FEM]. In sum, the Oromo diminutive manifests stacked *n*P s where the highest *n* feature is the gender of the whole nominal. Cross-linguistically, this pattern has also previously been reported for Somali (Kramer 2015), German (Wiltschko 2007), Dutch (Dressler & Barbaresi 1994) and Russian (Steriopolu 2008).

6. Conclusion

The present study addressed four objectives; (1) established the position of the Oromo gender system in Corbett’ (1991) typology of gender; (2) illustrated cross-dialectal gender assignment and agreement variation among the Oromo dialects; and (3) explained major factors that contribute to the gender variation among the Oromo dialects, and (4) described Oromo gender assignment and gender agreement based on the

framework of Distributed Morphology. Regarding the first objective, it was illustrated that there are two types of Oromo gender assignment systems; a form-based system which is proven to be a typical feature of the Eastern and Southern Oromo dialect, and a semantic-based system which is attributed to the Western and Northern Oromo dialects. It was further illustrated that gender in Oromo is non-alliterative. In Oromo, gender agreement targets verbs, different kinds of pronouns, and adjectives.

Regarding the gender variation, the study illustrated that gender assignment and agreement vary across the Oromo dialects; the Northern and Western dialects are alike, but the Eastern and Southern dialects are different both in terms of gender assignment and agreement. The major areas of difference include the agreement between adjectives and nouns, agreement between nouns and verbs and agreement between nouns and different kinds of pronouns. It was also illustrated that the historic *t* vs. *k* distinction has been completely neutralized in the Western and Northern dialects. The third objective addressed factors contributing to this variation. It has been argued that the phonology-based gender assignment in the Eastern and Southern Oromo dialects is the outcome of contact between the Oromo dialects and other Lowland East Cushitic languages, mainly Somali and Afar. The Western and Northern Oromo dialects have had contact with Semitic languages, mainly with Amharic. Due to this contact, gender assignment in these dialects shifted towards the default masculine Semitic gender system. It was also argued that the neutralization of the *t* vs. *k* distinction is the outcome of contact between Oromo and other Ethiopian languages, mainly the Semitic languages. These findings have some theoretical implications regarding diachronic gender loss. As noted by Schapper (2010), Enger (2004), Clamons (1992), Busterud et al. (2019), Igartua (2019) and Corbett (1991), contact among neighboring languages and varieties is the main source of gender loss. However, it seems that not all contact leads to gender loss. In the context of Oromo dialects, it is only the intense and durable contact that leads to gender loss, not just adult second language learning which was argued to be the case in Igartua (2019), Dahl (2004) and Trudgill (2013). The findings further imply that loss of gender assignment precedes the loss of gender agreement. More importantly, gender loss due to cross-dialectal influence seems more rapid and more influential than gender loss due to contact between languages. Regarding the fourth objective, it has been illustrated that Oromo gender can be analyzed using Kramer' (2015) morpho-syntactic approach. More importantly, it has been argued that, at least in languages that have a form-based gender assignment pattern, there is a predictable association between the root and the gender feature on *n*.

Finally, some words of caution are in order. The present study relied on a Speech Corpus that has about quarter of a million tokens. There is

also an asymmetry of data representation in the Corpus. Hence, we recommend similar future investigations, based on primary data.

References

- ACQUAVIVA, P. 2009. Roots and lexicality in distributed morphology. *York papers in linguistics*. eds. A. Galani, D. Redinger & N. Yeo, 1–21. York: University of York.
- ALPHER, B. 1987. Feminine as the unmarked grammatical gender: Buffalo girls are no fools. *Australian Journal of Linguistics* 7.2:169–187.
- APPLEYARD, D. 2015. Ethiopian Semitic and Cushitic: Ancient contact features in Ge'ez and Amharic. *Semitic languages in contact*. ed. B. Aaron, 16–32. Leiden: Brill.
- BANTI, G. 1988. Two Cushitic systems: Somali and Oromo nouns. *Autosegmental studies in pitch accent*. eds. H. Van der Hulst & N. Smith, 11–49. Dordrecht: Foris.
- BATES, E., DEVESCOVI, A., PIZZAMIGLIO, L., D'AMICO, S. & HERNANDEZ, A. 1995. Gender and lexical access in Italian. *Perception & Psychophysics* 57.6:847–862.
- BEYESSA, F. 2014. Major factors that affect grade 10 students' academic achievement in science education at Ilu Ababora General Secondary of Oromia Regional State, Ethi. *International Letters of Social and Humanistic Sciences* 21:118–34.
- BLAZEK, V. 2010. Glottochronological classification of Oromo dialects. *Lingua Posnaniensis* 52:2–27.
- BLENCH, R. M. 2013. Links between Cushitic, Omotic, Chadic and the position of Kujarge. *Proceedings of the 5th international conference of Cushitic and Omotic languages*. eds. S. Mari-Claude & M. van Hove, 67–80. Köln: Rüdiger Köppe.
- BULL, T. & SWAN, T. 2002. The representation of gender in Norwegian. *Gender across Languages* 2:219–249.
- BUSTERUD, G., LOHNDAL, T., RODINA, Y. & WESTERGAARD, M. 2019. The loss of feminine gender in Norwegian: A dialect comparison. *The Journal of Comparative Germanic Linguistics* 22.2:141–167.
- CARSTENS, V. 2010. Implications of grammatical gender for the theory of uninterpretable features. *Exploring crash-proof grammars*. ed. P. Michael, 33–57. Amsterdam: Benjamins.
- CHOMSKY, N. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- CLAMONS, R. 1992. *Gender in Oromo*. University of Minnesota. PhD diss.
- CLAMONS, C. 1993. Gender Assignment in Oromo. *Iverson principles and prediction: The analysis of natural language*. Eds. E. Mushira & K. Gregory K., 269–284. Amsterdam: John Benjamins.
- CLAPHAM, C. 1988. *Transformation and continuity in revolutionary Ethiopia*. Cambridge: Cambridge University Press.
- CORBETT, G. 1991. *Gender*. Cambridge: Cambridge University Press.
- CRASS, J. & MEYER, R. 2011. Ethiosemitic-Cushitic Language Contact. *The Semitic languages: An international handbook*. ed. W. Stefan, 1266–1275. Berlin: De Gruyter Mouton.
- DAHL, Ö. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.
- DEKEYSER, X. 1980. The diachrony of the gender systems in English and Dutch. *Historical Morphology* 17:97–111.
- DEMIE, B. 2010. *The impact of Afaan Oromo dialectal variations on teaching-learning process*. Addis Ababa University: Addis Ababa. PhD diss.

- DUKE, J. 2010. Gender reduction and loss in Germanic: Three case studies. *Kontrastive germanistische linguistik*. eds. A. Damaris & K. Sebastian, 643–673. Hildesheim: Olms.
- DRESSLER, U. & BARBARESI, M. 1994. *Morphopragmatics: Diminutives and intensifiers in Italian, German, and other languages*. Berlin, New York: Mouton de Gruyter.
- ENGER, O. 2004. On the relation between gender and declension: A diachronic perspective from Norwegian. *Studies in Language* 28:51–82.
- FELEKE, T. L. 2021. Ethiosemitic languages: Classifications and classification determinants. *Ampersand* 8:100074.
- FELEKE, T. L. (2023). Issues in classifying and mapping the Semitic languages of Ethiopia. *Language, society and the state in a changing world*. eds. S. Brunn & R. Kehrein, 33–85. Cham: Springer International Publishing.
- GASSER, M. 2010. A dependency grammar for Amharic. In *proceedings of the workshop on language resources and human language technologies for Semitic Languages*, Valletta, Malta.
- GEMEDA, H. 2006. *Functional categories in Afan Oromo: A minimalist approach*. Addis Ababa University. MA thesis.
- HARRIS, J. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22:27–62
- HASSEN, M. 2015. *The Oromo and the Christian kingdom of Ethiopia: 1300–1700*. Suffolk: Boydell & Brewer.
- HAWINE, A. 2009. *The structures of nominal clauses in Afan Oromo a minimalist approach*. Addis Ababa University. PhD diss.
- HAYWARD, J., & CORBETT, G. 1988. Resolution rules in Qafar. *Linguistics* 26.2:259–280.
- HEINE, B. 2012. The Waata Dialect of Oromo: Grammatical Sketch and Vocabulary, Language and Dialect Atlas of Kenya. *Journal of the International African Institute* 55.2:228–232.
- HETZRON, R. 1980. The limits of Cushitic. *Sprache und Geschichte in Afrika* 2:7–125.
- HOCKETT, C. F. 1958. *A course in modern linguistics*. New York: Macmillan.
- HOPP, H. 2013. Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research* 29.1:33–56.
- IGARTUA, I. 2019. Loss of grammatical gender and language contact. *Diachronica* 36.2:181–221.
- KAYNE, R. S. 2005. On parameters and on principles of pronunciation. *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*. eds. H. Broekhuis, N. Corver, R. Huijbregts, U. Kleinhenz & J. Koster, 289–299, Mouton: De Gruyter Mouton.
- KEBEDE, K. 2012. *Towards the genetic classification of the Afaan Oromoo dialects (V.2)*. Rudiger Koppe Verlag. Kologn: Germany
- KEFALE, A. 2010. Federal restructuring in Ethiopia: Renegotiating identity and borders along the Oromo–Somali ethnic frontiers. *Development and Change* 41.4:615–635.
- KRAMER, R. 2009. *Definite markers, phi features and agreement: A morphosyntactic investigation of the Amharic DP*. University of California, Santa Cruz. PhD diss.
- KRAMER, R. 2015. *The morphosyntax of gender* (Vol. 58). Oxford: Oxford University Press.
- LESLAU, W. 1945. The influence of Cushitic on the Semitic languages of Ethiopia a problem of substratum. *Word* 1.1:59–82.
- LESLAU, W. 1995. *Reference grammar of Amharic*. Wiesbaden: Harrassowitz Verlag.

- LES LAU, W. 2000. *Introductory grammar of Amharic* (Vol. 21). Wiesbaden: Harrassowitz Verlag.
- LOHNDAL, T. & WESTERGAARD, M. 2016. Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology* 7:344.
- MARCHAL, H., BIANCO, M., DESSUS, P. & LEMAIRE, B. 2007. The development of lexical knowledge: Toward a model of the acquisition of lexical gender in French. *Proceedings of the second European cognitive science conference*. eds. S. Vosniadou, D. Kayser & A. Protopapas, 268–273. Mahwah: Erlbaum.
- MARCUS, H. G. 2002. *A history of Ethiopia*. California: University of California Press.
- MATASOVIĆ, R. 2004. *Gender in Indo-European*. Heidelberg: Winter.
- MEKONEN, H. 2002. *Lexical standardization in Oromo*. Addis Ababa University. MA thesis.
- MEWIS, C. J. 2001. *A grammatical sketch of written Oromo*. Koln: Rudiger Koppe Verlag.
- MOHAMMED, A. & ANDRZEJ, Z. 1990. *Handbook of the Oromo Language*. Stuttgart: Steiner.
- MOUS, M. 2008. Number as an exponent of gender. *Interaction of morphology and syntax: Case studies in Afro-asiatic*. eds. S. Erin & F. Zygmunt, 137–160. Amsterdam: Benjamins.
- MOUS, M. 2012. Cushitic. *The Afro-asiatic languages*. eds. F. Zygmunt & S. Erin, 342–422. Cambridge: Cambridge University Press.
- NEGESSE, F. 2015. Classification of Oromo dialects: A computational approach. *International Journal of Computational Linguistics* 6.1:1–10.
- OLANI, W. 2017. *Inflectional morphology in Oromo*. Addis Ababa University. MA thesis.
- OOMEN, A. 1981. Gender and plurality in Rendille. *Afro-asiatic Linguistics Los Angeles* 8.1:35–75.
- OWENS, J. 1985. *A grammar of Harar Oromo (Northeastern Ethiopia)* (Vol. 4). Hamburg: Buske.
- PICALLO, M. 1991. Nominals and nominalization in Catalan. *Probus* 3:279–316.
- RALLI, A. 2002. The role of morphology in gender determination: evidence from Modern Greek. *Linguistics* 40, 519–51.
- RIENTE, L. 2003. Ladies first: the pivotal role of gender in the Italian nominal inflection system. *McGill Working Papers in Linguistics* 17.2:1–54.
- ROCA, I. M. 1989. The organization of grammatical gender. *Transactions of the Philological Society* 87:1–32.
- RITTER, E. 1993. Where's gender? *Linguistic Inquiry* 24:795–803.
- RODINA, Y. & WESTERGAARD, M. 2015. Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics* 27.2:145–187.
- SAEED, J. 1999. *Somali*. Amsterdam: Benjamins.
- SCHAPPER, A. 2010. Neuter gender in eastern Indonesia. *Oceanic Linguistics* 49.2:407–435.
- STERIOPOLO, O. 2008. *Form and function of expressive morphology: A case study of Russian*. University of British Columbia. PhD diss.
- TAYE, B. A. 2019. The medium of instruction in Ethiopian higher education institutions: Kotebe Metropolitan University Case study. *African Journal of Teacher Education* 8:1–24.
- TEGEGNE, W. 2015. The inclusion of dialects in education: An exploration into the use of Afan Oromo dialects in primary educational context. *International Journal of Language and Linguistics* 3.6:360–366.

- TEMAM, A. 2020. Gender in Afaan Oromo (Oromoic). Unpublished research paper, received via ResearchGate.
- TOSCO, M. 2000. Cushitic overview. *Journal of Ethiopian Studies* 33.2:87–121.
- TRUDGILL, P. 2013. Gender maintenance and loss in Totenmälet, English, and other major Germanic varieties. *Search of universal grammar: From Old Norse to Zoque*. ed. T. Lohndal, 77–108. Amsterdam: John Benjamins Publishing.
- TSEGAYE, M. T. 2017. *Plural gender: Behavioral evidence for plural as a value of Cushitic gender with reference to Konso*. Leiden University. PhD diss.
- WAKWEYA, G. 2017. Inflectional morphology in Mecha Oromo. *Journal of Languages and Culture* 8.8:110–140.
- WILTSCHKO, M. 2007. Why should diminutives count. *Organizing grammar: studies in honor of Henk van Riemsdijk*. eds. B. Hans, C. Norbert, R. Huybregts, K. Ursula & K. Jan, 669–678. Berlin: Mouton de Gruyter.
- WOLDEMARIAM, H. & LANZA, E. 2014. Language contact, agency and power in the linguistic landscape of two regional capitals of Ethiopia. *International Journal of the Sociology of Language* 228:79–103.
- WORKNEH, B. 2011. *DP internal agreement in Amharic: A reverse agree solution*. University of Tromsø. MA thesis.
- YADATE, A. R. 2019. *Linguistic sexism in gender assignment systems of Afan Oromo*, Amharic and Gamo. Selinus University of Sciences and Literature, Italy. PhD diss.
- YIMAM, B. 1986. *The phrase structures of Ethiopian Oromo*. School of Oriental and African Studies. PhD diss.
- ZABORSKI, A. 1991. Ethiopian language subareas. *Unwritten testimonies of the African past*. Ed. P. Stanis & R. Eugeniusz, 123–134. Warsaw: Wydawnictwa Uniwersytetu Warszawskiego.

Received July 1, 2021

Accepted April 28, 2023

Tekabe Legesse Feleke

AcqVA Aurora Research Center, Department
of Language and Culture

The Arctic University of Norway

Tromsø

Norway

tekabel.feleke@uit.no

Terje Lohndal

AcqVA Aurora Research Center, Department
of Language and Culture

The Arctic University of Norway

Tromsø

Norway

terje.lohndal@ntnu.no

Department of Language and Literature
Norwegian University of Science and

Technology

Trondheim

Norway