

---

---

## **Linguistic Variation within Sidaama**

Fekede Menuta<sup>1</sup> and Kebede Teklemichael<sup>2</sup>

### **Abstract**

Sidaama, which also designates an ethnic group that speaks the language, is a member of the Highland East Cushitic languages of Ethiopia. So far there has been no empirical study regarding linguistic variation within Sidaama. This article provides descriptive account of linguistic diversity within Sidaama. The study used both qualitative and quantitative data to answer the research questions. The data for the study were collected from four sample areas; namely, Harbegona, Dara, Shebedino and Lokkaabaya districts. A total of 32 native speakers of Sidamuafoo, eight from each sample area, were used as key informants. A total of 205 lexical items were elicited following Swadesh word lists as a guide. These lexical items were phonemically transcribed and then analyzed using Cog1.3.1 lexicostatistics software to determine the linguistic similarity and difference across the sampled areas. Thematic comparison was also made to trace other linguistic similarities. Both UPGMA and neighbor joining<sup>3</sup> analysis showed that the language exhibits between 80%-88% similarity across the varieties. This implies that the language is significantly homogeneous, and intelligibility level is relatively high, with the risk of only 12%-20% of misunderstanding. However, this same analysis showed that there are two possible dialect areas: Dara-Harbegona on one side and Shebedino-Lokkaabaya on the other. In addition to the lexical variations, there are a few morphological areal differences. It is the researchers' presumption that the current mother tongue education program, the Sidaama language use in media, and the language standardization efforts made may harmonize the variation, and reduce the existing areal differences.

**Key Words:** Cushitic, Language, Lexicostatistics, Sidaama, Regional-variation

---

<sup>1</sup> Fekede Menuta is associate professor of Linguistics and Communication in the Department of Language and Literature of Hawassa University, Ethiopia. E-mail: [mnutafekede2012@gmail.com](mailto:mnutafekede2012@gmail.com)

<sup>2</sup> Kebede Teklemichael is assistant professor of Linguistics and Communication in the Department of Language and Literature of Hawassa University, Ethiopia. E-mail: [micheal.kebede19@gmail.com](mailto:micheal.kebede19@gmail.com)

<sup>3</sup> UPGMA is unweighted pair group method with arithmetic mean, and Neighbor-joining is a bottom-up clustering method for the creation of phylogenetic or evolutionary tree.

## 1. Introduction

### 1.1 The Sidaama People and their Language

While the term ‘*Sidaama*’ refers to the people, the term ‘*Sidaamu afoo*’, which literally means ‘the mouth of Sidaama’ is used to denote the language. In some literature, we find the term ‘*Sidamo*’, referring to both the people and the area that the people live in. However, this term was used mainly to refer to a province (‘*Kifle Hager*’, literally means ‘part of a country’ that covered Sidaama and other HEC language speaking groups (cf. (Anbessa 2000:12). According to some oral narratives, the name ‘Sidaama’ etymologically originates from the tribe name ‘Siddo’. This tribe fought and defeated a tribal group called “Hoofa”, who probably are Omotic Speakers (cf. Dukamo, 2014:1-2; Gasparini, 1983: VI; Beetana,1983:83). According to this story, the Hoofa, being astonished by the sudden attack and subsequent defeat, said the word ‘Siddo-ma’ (lit. ‘Siddo’, the tribe name and ‘-ma’, ‘what’), to imply “What a powerful tribe.” Over a period of time, Siddoma has changed to Sidaama by degeminating /d/ and by adding the vowel /o/ to compensate the degemination, which in turn is harmonized with the vowel of suffix {-ma}. This same oral story was mentioned by informants from the Harbegon district. There is also claim that Sidaam for Oromo people refers to non-Oromo; hence, is less specific. The exact etymology of the term is, therefore, open for the future ethno-historical research. We will use Sidaama in this article to refer to both the ethnic group and the language.

The Sidaama people live in the Sidaama Reginal State. Hawassa, which currently is serving as the capital city of the Sidaama, and the SNNPRS regional states, is found 275 KM to the south east of Addis Ababa. The Region is being restructured into four Zones following the recognition of Sidaama as the 10<sup>th</sup> reginal state of Ethiopia following the November 23, 2019 referendum. We have presented here the districts prior to the restructuring. There were 19 districts; namely, Hawassa Zuria, Wondo

Genet, Borricha, Malga, **Shebedino**, Gorche, Daale, Wonsho, Bursa, **Lokkaabaya**, Alettawondo, **Harbegona**, Hula, **Dara**, Chukko, Cirre, Haroressa, Bansa, and Booni. There are also two important reform towns, namely Aletta and Yirgalem. The present study has purposively selected four of the districts which are written in bold here. Shebedino was selected as a central dialect; Lokkaabaya was chosen it is contact area with Wolayta hence there might be linguistic infiltrating; Harbegona was chosen because it has contact with Oromo and the Hadicho of Sidaama who socially said to differ are found there; Dara was chosen as a contact area with Gedeo.

Majority of the Sidaama people are farmers who make their living by producing different crops and rearing animals. Ensete (*Ensete ventricosum*) is the main crop plant out of which they produce their staple food weese. Coffee is an important cash crop for the people. They breed animals such as cows, sheep, goats, etc. Fishing is also an important food source for the people living around Lake Hawassa.

Sidaamuafo belongs to the Highland East Cushitic (HEC) branches of the Cushitic languages phylum of the Afro-asiatic. There have been several classifications of Cushitic languages in general and HEC languages in particular. (cf. Hudson 1976: 244, 2007:529; Tosco 2000: 90) We have adopted the classification by Tosco 2009 shown in Figure-1:

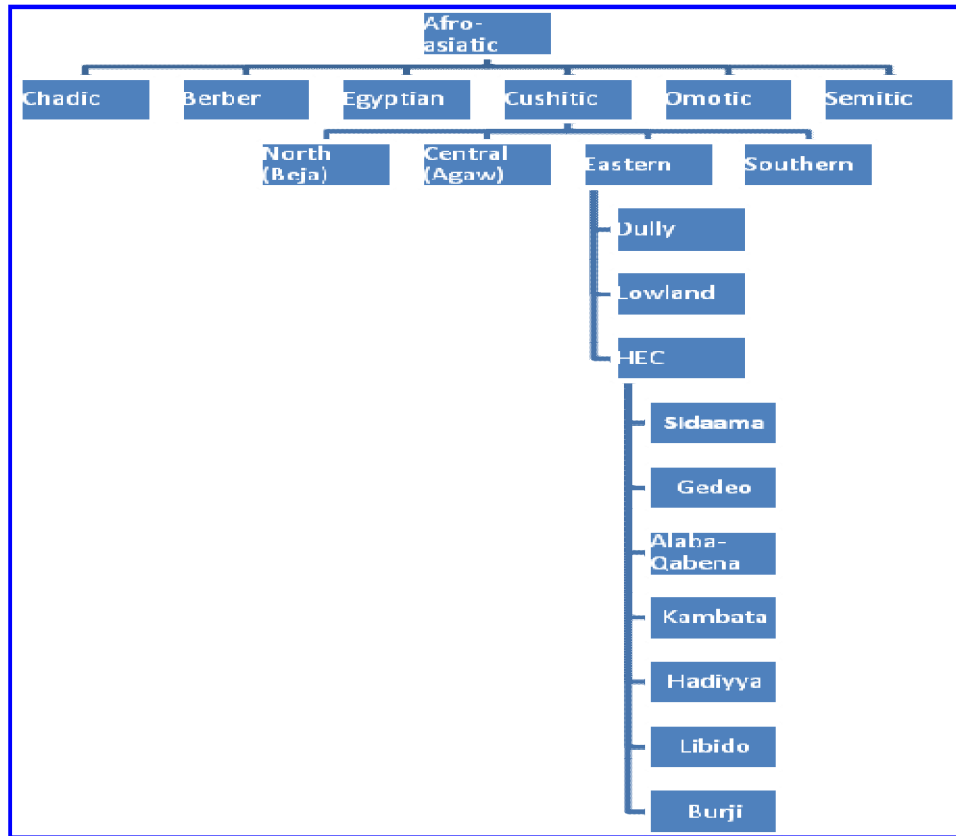


Figure-1: Position of Sidama within HEC (Tosco, 2009:12)

Some works, provide an internal grouping of HEC languages. Hudson (2007:529) for instance groups Burji as relatively divergent compared to other HEC languages. Sidaama and Gedeo are grouped as relatively closer languages. Kambata, Tembaro and Qabena are also identified as closely related varieties. There is another language variety called Halabissa (i.e. the language of Halaba people), which is not actually included in this group by Hudson (2007) but which we think should be a part (cf. Fekede, 2012). Hadiyya and Libido (also called Mareqo) are closely related language varieties.



## 1.2 Problem Statement

Language variation in Sidaama is one of the least investigated aspects of the language. Though some scholars have already noticed both social and geographical variations, there are no empirical studies that show levels of linguistic variations in Sidaama. The social variation has been reported by Anbessa (1987) in an article titled *Balliffa: Women's Speech among the Sidaama*. In this article, he states that there is a considerable difference between the speech of men and women. The language variation is triggered by taboos related to names of in-laws, spouses and sub-clans (Anbessa 1987:44). Anbessa (1987) reports that a similar gender-based and taboo driven language variations are witnessed in Hadiyya and Arisi Oromo. Regarding regional variation in Sidaama, Kawachi (2007:3) has the following to say:

As far as I know, there does not seem to be any dialectal study on Sidaama. Hudson (1976: 233) states that Sidaama is “little differentiated into dialects” and that “speakers of the different centers and Sidamo clans are identifiable by speech, but the differences are minimal”. However, according to my consultant, there are large differences among both regional and social dialects of Sidaama. In particular, there are two clans whose varieties of Sidaama differ significantly from those spoken in the other five or six clans. One of them is the *Yanase* clan in the north, which is geographically separated from other clans. The other is the *Haadiicfffo* (Brøgger 1986: 34), whose members are considered to do potters' jobs and seem to be despicably treated by members of other clans. Dialectal differences of Sidaama await further investigation.

The present study tries to uncover the extent of regional variation in different areas of Sidaama to fill in the gap that Kawachi had reported awaiting further study. We, however, do not opt to make a study on social variations. It is also worth to report that this study is a part of a larger

undergoing project that aims to compare Diversity and Shared Values among five Cushitic Languages speaking groups.

### **1.3 Objective**

The main objective of the study is to provide a descriptive account of linguistic variation and similarity within Sidaama. It has the following specific objectives:

- a) To find out the level of lexical and Phonetic similarity and difference
- b) To estimate the communication cost across the varieties
- c) To uncover morphological variations if any

### **1.4 Literature review**

Linguistic variation includes phonological, lexical, syntactic and semantic. The phonological variation focuses on sounds and phonological rules; the lexical variation focuses on difference in lexical items which can be viewed synchronic or diachronically and may include the simple and complex words with affixes; syntactic variation is about the ways words are structured in compared languages or language varieties; and semantic variation is concerned with how meaning varies synchronically due to sociolinguistic variables and diachronic semantic changes (cf. Bayley and Lucas, 2007).

Language variation is studied from two broader perspectives or focuses; namely, focus on users and focus on uses. Based on focus on users, we have regional or spatial dialects and social dialects. The social dialects consider variables, such as age, gender, and ethnicity. Based on focus on uses, we use variation due to style, context, register, politeness, stereotyping, cognition and attitude (cf. Holmes, 2013). The source of data for sociolinguistics study can be large survey, interview, observation, corpora, elicited lexical items, and other sources, such as census report (Schilling-Estes, 2007:165-175). The analysis can be qualitative, quantitative (cf.

Tagliamonte, 2007:190) or mixed- both qualitative and quantitative methods used in a single research.

This article is on phonological and lexical variation in Sidaama. It focuses on users than the uses, and focuses particularly on dialect variation than social dialect variation. The data collection is based on elicitation of lexical items, and the analysis is mixed methods as discussed in (1.5).

There are several linguistic descriptions in Sidaama. The most common ones are Dukamo (2014) 'Documentation of the form and function of Sidaama demonstratives'. He identifies four ways of contrast for nominal demonstrative. These are proximal (kuni (SGM), tini(SGF) and kuri(PL)), mesioproximal (hakku (SGM), hatti(SGF) and hakkuri(PL), mesiodistal (kuu'u(sgm), tii'i(SGF) and kuu'uri(PL), and distal (ku''u(SGM), ti''i(SGF) and ku''uri(PL). Dukamo (2014) also provides four nominal adverbs that mark location differences from speaker's point of reference. These locative adverbs include: kaw-iichi (here (PRX), hakk-iichi (there (MPRX), kaa'-iichi (there after (MDST) and ka''-iichi (over there (DST)). He also provides four accusative forms, each having three allomorphs, parallel with the four demonstrative adverbs mentioned above (cf, Dukama, 2014:31). Anbessa (2000) a PhD thesis titled A Grammar of Sidaama provides the grammatical description of Sidaama In his article titled 'Balliffa: Women's Speech among the Sidaama', Anbessa (1987) identifies gender-based social variations in language by women. He provides a list of lexicons women use as substitutes for linguistic items that men or the society at large use while dealing with taboos. Kawachi (2007) in 'A Grammar of Sidaama (Sidamo): A Cushitic Language of Ethiopia' presents a complete grammatical description of Sidaama. Kawachi (2007) in an article titled 'Feelings in Sidaama' identifies verbs that are used to express feelings. Eshetu (2010) identified sound correspondence and difference within HEC languages in which Sidaama is one by conducting a comparative study of Burji, Hadiyya, Kambaata, Sidaama and Gedeo. In 'Generating Narratives: Interrelations of Knowledge, Text Variants, and Cushitic Focus Strategies', Wedekind (1990) described various communicative strategies among selected Cushitic

languages. He also compared and contrasted the main phonological features of three HEC languages in his 1980 work titled 'Sidamo, Gedeo (Derasa), Burji: Phonological Differences and Likenesses'.

## 1.5 Methodology

### 1.5.1 Research Design and Method

The study followed a cross-sectional descriptive design and mixed methods. We used comparisons of lexical and phonetic distance using statistical package called Cog to find out dialect variation and similarities. The specific methods used were interview and elicitation of lexical items included in Swadesh lists and some more items which are commonly and frequently used. A total of 205 lexical items have been compared. The words were elicited from informants who live in Lokkaabaya, Shebedino, Harbegona and Dara, which are purposefully selected districts within the Sidaama Zone to represent linguistic contact areas, and a central dialect which is relatively stable. (cf. sampling 1.5.2). We have also compared morphological variations with narrative texts.

The lexical items were transcribed phonemically using IPA, and then a comparison of lexical similarity was calculated using the software *cog* developed by SIL. Cog assigns values for sound correspondences. It uses the rule of deletion or insertion for sounds that are missing in one variety but occur in the other. Exact correspondences receive the same value, but differences are computed based on the number of shared features. The default for calculating difference is a single feature. To be specific, if two sounds of two dialects do not differ in any feature, they receive the same value; if they differ in one feature, they receive the next highest similarity value; the more the difference in feature, the higher the difference.

The Cog software provides different outputs based on what the researcher wants to know. It can provide percentages of shared lexicons between two dialects, percentages of shared phonetic features, and two types of

## Linguistic Variation within Sidaama - Fekede and Kebede

dendrograms-*UPGMA* and *Neighbor-joining*. The *UPGMA* method has extensively been used in bioinformatics and in comparative linguistics to construct a rooted tree based on a pairwise similarity matrix. It assumes a constant rate of evolution; this has been considered its weakness by some scholars. Neighbor-joining analysis however produces output without considering the rate of evolution. Thus, it is up to the researcher to choose the one that fits his/her research needs.

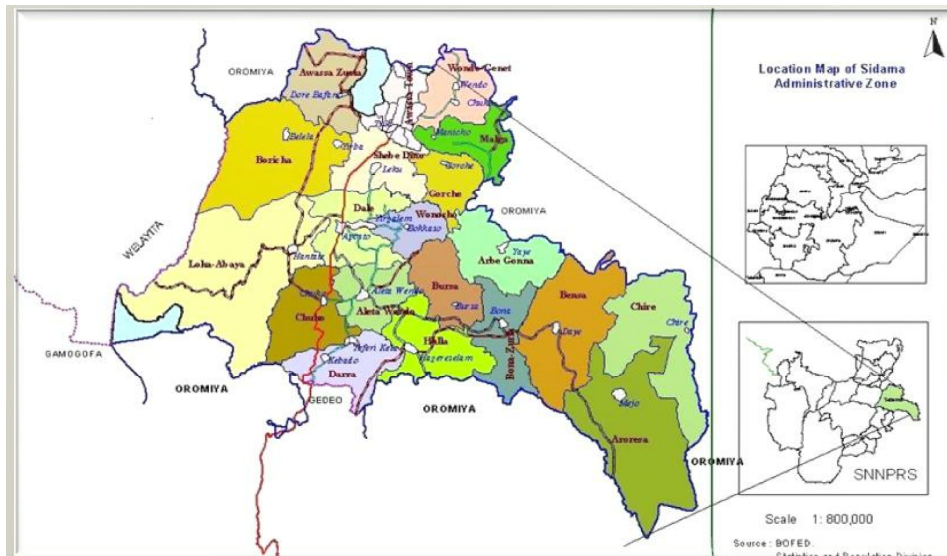


Figure 1: Administrative map of Sidaama Regional State

### 1.5.2 Sampling

#### a) Area Sampling

As we had no information with regard to language variation in Sidaama, we have consulted experts in the language and culture office of the Sidaama Regional State at Hawassa. They informed us about the potential areas for variation, because of language contact situations, and areas with limited chance for diversity. Based on this suggestion, we used purposive sampling, and chose *Lokkaabaya*, *Darra*, *Harbegona* and *Shebedino*. *Lokkaabaya* borders with Welayta in the west, Oromo in the south, and Boricha, Dale, and Chuko districts of Sidaama in the north, northeast, and east,

respectively. *Darra* borders with Gedeo Zone in the South, Oromia in the southeast, Hulla, Aleta Wendo and Chuko districts of Sidaama in the east, north and North West, respectively. Harbegona largely borders with Oromiya in the east, Gorche, Wonsho, Bursa, Bona-Zuria, Bensa districts of Sidaama in the north, northwest, west, south, and Southeast. Shebedino is the central area completely encircled by other districts of Sidaama; namely, Goriche, Dale, Boricha, Awassa Zuria, Awassa town, and Malga in the East, South, West, North West, north, and northeast, respectively. This sampling allowed language contact areas with possible linguistic infiltration, such as Lokkaabaya, Darra, Harbegona and relatively central dialect which has no contact with non-Sidaama languages of the area, such as Oromo, Gedeo and Wolayta. This kind of purposive or judgment sampling is a preferred method in sociolinguistics than random sampling (cf. Schilling-Estes, 2007:169).

#### **b) Participant sampling**

We purposefully selected 8 key informants (4 males and 4 females) in each of the four areas, hence, we had a total of 32 participants. The criteria for selection were inclusion of both male and female, age gaps (youngsters and elders) and people from different villages of the same district with different social status. We use these key informants exclusively for linguistic elicitation. We also had a survey on language attitude and knowledge of other Cushitic languages. We used 385 people for this survey. However, the result of the survey is not included in this article because it is meant for comparing across Cushitic groups than within Sidaama.

## 2. Presentation of Results

### 2.1 Lexical and Phonetic similarity comparison

#### 2.1.1 Lexical comparison

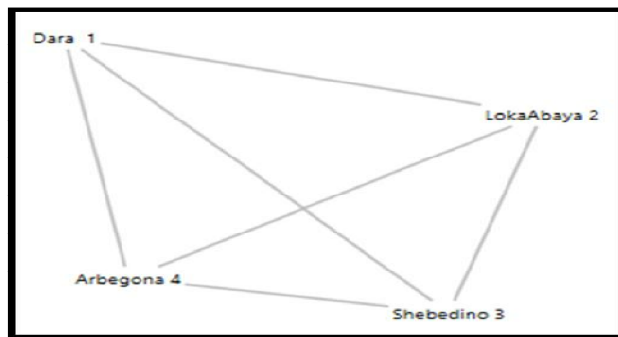
Figure-1 Below shows the results of the lexical comparison

	Dara 1	Arbegona 4	Shebedino 3	LokaAbaya 2
Dara 1		86	80	80
Arbegona 4	86		87	82
Shebedino 3	80	87		88
LokaAbaya 2	80	82	88	

The highest lexical similarity is observed between Shebedino and Lokkaabaya (88%), and then between Shebedino and Harbegona (87%). The third and fourth similarity is between Harbegona and Dara (86%) and between Harbegona and Lokkaabaya (82%), respectively. Dara has a relatively lowest similarity with Shebedino and Lokkaabaya (80%).

*Figure-1a: Lexical similarity matrix*

The network graph below shows lexical similarity (fig. 1b):



*Figure-1b: Lexical similarity network*

The lexical similarity network demonstrates that Dara is slightly diverging from the rest, (i.e. from Harbegona, Shebedino and Lokkaabaya. Relatively, Harbegona and Shebedino are closer to each other than the rest, though Shebedino and Lokkaabaya seem to have the highest lexical similarity. Dara seems to be more distant to Shebedino than it is to the other two: Harbegona and Lokkaabaya.

2.1.2 Phonetic comparison

The situation is slightly different when it comes to the phonetic similarity. All the four areas seem to have similar phonetic sounds, except for the Dara variety which is slightly divergent, as it is demonstrated with the phonetic similarity matrix in Figure 2a and the network graph in 2b.

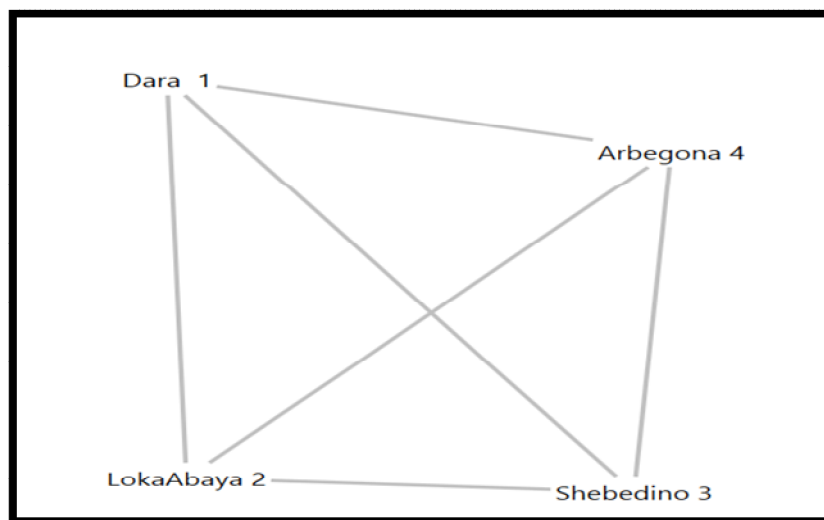
Similarity Matrix

	Dara 1	Arbegona 4	Shebedino 3	LokaAbaya 2
Dara 1		88	84	86
Arbegona 4	88		88	86
Shebedino 3	84	88		90
LokaAbaya 2	86	86	90	

Figure 2a: Phonetic similarity matrix

The phonetic similarity matrix shows that the phonetic sounds among the four areas of Sidaama are nearly the same. The maximum difference observed is 16% between Shebedino and Darra while the maximum similarity is 90% between Lokkaabaya and Shebedino.





*Figure-2b: Phonetic similarity network graph*

In most of the cases, the graph is relatively leveled, implying uniformity across the four areas, especially between Lokkaabaya and Shebedino and between Shebedino and Harbegona. The smallest phonetic distance is observed between Lokkaabaya and Shebedino. The phonetic distance between Dara and Lokkaabaya is as similar as between Shebedino and Harbegona.

## **2.2 Classification of the areal variations:**

We tried to classify the ‘varieties’ based on UPGMA analysis, which helps to determine evolutionary change, and Neighbor-joining analysis, which allows to identify some form of grouping without considering rate of evolutionary change. Though these two analyses often produce different classifications in languages whose time of separation is longer, this is not the case with regard to Sidaama.

### **a) UPGMA Analysis**

We made two outputs of classification based on whether evolutionary changes are assumed (UPGMA) or not, Neighbor joining. Though these two

analyses often produce different classification in some languages whose separation time is longer, the outputs are the same with regard to Sidaama. The UPGMA implies that the language did not diverge in the remote past. Below is UPGMA (rooted) analysis of the four areas:

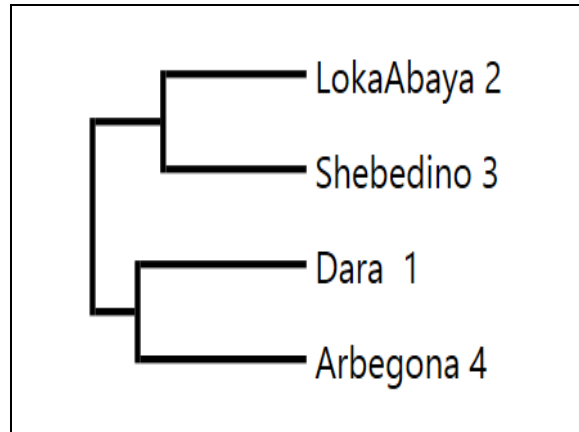


Figure 3a: Dendrogram of UPGMA (rooted Analysis)

### b) Neighbor-joining Analysis

To group the language varieties sampled in the study into their closest language variety, we have run *neighbor joining* analysis and we found the result shown in figure-3:

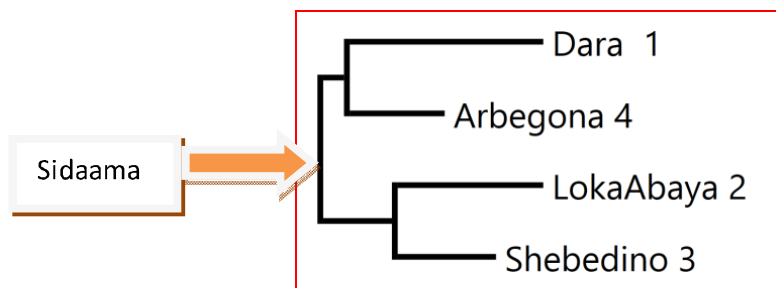


Figure-3b: Dendrogram of neighbor joining

The neighbor joining analysis showed that there are relatively two groups branching out from Sidaamuafo: Harbegona-Dara and Shebedino-Lokkaabaya. It further shows that Harbegona variety slowly diverged to the Dara-Harbegona, and Shebedino variety to Lokkaabaya Shebedino.

### 2.3 Communication Cost Analysis

Communication cost analysis (CCA) is concerned with the analysis of level of communication difficulty against the possible level of intelligibility. The highest the intelligibility among languages or language variety groups, the lowest the communication cost, and vice versa. A perfect communication or intelligibility is based on 100% shared vocabulary, which is possible among speakers of a language. Thus, we calculated the communication cost for the four Sidaama language varieties by subtracting the shared vocabulary percent value from 100%, the perfect communication level. Table-1 below shows the result.

<b>Language Varieties</b>	<b>100 -Shared voc.</b>	<b>Cost</b>	<b>Rank</b>
Shebedino - Lokkaabaya	100-88	12	1
Shebedino – Harbegona	100-87	13	2
Harbegona – Dara	100-86	14	3
Shebedino – Dara	100-80	20	4
Lokkaabaya -Dara	100-80	20	4

*Table-1: Communication cost analysis*

Most researches on intelligibility consider 75% shared vocabulary (i.e. 25% communication cost) allows reasonably good communication among speakers of language varieties (cf. Fekede, 2015; Casad, 1974). Of course, intelligibility below 75 but above 61 can still allow fair communication. As shown in the table, the minimum shared vocabulary is 80% and the maximum communication cost is 20%. This implies that intergroup communication is possible without much difficulty. In other words, speakers of Sidaamuafo constitute a mutually intelligible linguistic community with relatively low communication difficulty.

Communication between Shebedino and Lokkaabaya costs the least (12% difficulty), followed by Shebedino-Harbegona, with 13% cost. Harbegona and Dara are the third with regard to smooth intelligibility. of course, with a communication cost of 14%. Both Shebedino-Dara and Lokkaabaya-Dara relatively have the highest communication cost, both 20%. Based on these, we can conclude that Shebedino is relatively the most shared central language variety while Dara is the least shared and is a peripheral variety.

## 2.4 Morphology

### Pronoun

The pronouns of Sidaamuafu distinguish singular and plural in first, second and third person. However, gender distinction in pronouns is made only in third person singular. Table-2 shows the pronoun system in the nominative case form.

Person	Dara	Lokkaabaya	Shebedino	Harbegona
1SG	an-i	an-i	an-i	an-i
1PL	niʔke	niʔke	niʔnke	niʔnke
2SGMF	at-i	at-i	at-i	at-i
2PL	kiʔne	kiʔne	kiʔne	kiʔne
3SGM	ʔis-i	ʔis-i	ʔis-i	ʔis-i
3SGF	ʔis-e	ʔis-e	ʔis-e	ʔis-e
3PL	ʔinsa	ʔinsa	ʔinsa	ʔinsa

*Table-2: Independent pronouns of Sidaamuafu in nominative case*

The pronouns are relatively the same across the varieties. However, small variation is witnessed in 1PL. For example, Dara and Lokkaabaya have niʔke while Shebedino and Harbegona have niʔnke with -n- inserted between glottal stop /ʔ/ and velar stop /k/ to mark first person plural. Table 3 shows the person-number-gender markers:

	Person	Number	Gender
1SG	an-	-ø-	-i
1PL	ni-	-k-	-e
2SG	at-	-ø-	-i
2PL	ki-	-ʔ(n)-	-e
3GM	ʔis-	-ø-	-i
3SGF	ʔis-	- ø-	-e
3PL	ʔis-	-na	-ø

Table-3: Person, number and gender markers of nominative pronoun

Gender is distinguished only in third person singular, where masculine is marked with -i and feminine with -e. The 1SG and 2SG do not distinguish gender but formally assume the default masculine gender. Gender is completely neutralized to feminine in plural nouns and pronouns in Sidaama, as it is the case in most HEC languages. The 1PL and 2PL take {-e} for feminine but 3PL ends with {-a}. Note that 3PL is *ʔis-na* underlying, and it becomes *ʔinsa* due to the phonological process of metathesis, *sn-ton*.

### Locative Demonstrative Adverbs (LDA)

Unlike Dukamo’s (2014) findings, we have got only three ways of distinctions for locative adverbs 'here', 'there', and ‘over there’, which correspond proximal, medial and distal positions from the speaker. The following table shows the LDA in the four Sidaama areas:

Gloss	Dara	Shebedino	Lokka	Harbegona
‘Here’	ko-ʔe	ko-ʔe	ko-ʔe	ko-nne
‘There’	ko-ʔo	ko-ʔo	ko-ʔo	ha-kko
‘Over there’	ko-oʔo	ko-oʔo	ko-oʔo	ko-oʔo

Table-4: Locative demonstrative adverbs

Regarding demonstratives, all the three areas show similarity; however, Harbegona demonstrates a slight variation for ‘here’ and ‘there’. The locative marker of ‘here’ is {-ʔe} in Darra, Shebedino and Lokkaabaya, but {-nne} in

the case of Harbegona. Similarly, the locative marker for 'there' is {-ʔo} in Darra, Shebedino and Lokkaabaya, but it is {-kko} in Harbegona. What is more, the basic form, which is 'ko- 'elsewhere becomes 'ha- 'in the case of Harbegona.

### Noun Inflection

#### Number

Sidaamuafɔ distinguishes singulative, collective and plural nouns. The collective is marked by '{-(tʃ)}', which itself has allomorphic variations.

SG	Gloss	Plural	Gloss
man- <i>fu</i>	'a man'	Manna	'men'
man- <i>fɔ-te</i>	'a woman'	meen-to	'women'
saa	'a cow/'cow'	saa- <i>da</i>	'cows'
beedda- <i>kko</i>	'a star'	beedda- <i>he</i>	'stars'
barra	'day'	barru- <i>wa</i>	'days'
wadʒdʒo	'white'	wadʒdʒu- <i>de</i>	'white ones'
kolli- <i>fɔ</i>	'black'	kolli- <i>dda</i>	'black ones'
duumo	'red'	duumma- <i>dda</i>	'red ones'
jema- <i>ko</i>	'mouse'	jema- <i>he</i>	'mice'
doobbi- <i>tʃo</i>	'lion'	doobb- <i>e</i>	'lions'
amboom- <i>i</i>	'hyena'	amboom- <i>a</i>	'hyenas'

Table-5: Number in nouns

We have distinguished three variants of singulative in Sidaama; i.e. -(k)ko; -(f)fo, and -(tʃ)tʃo. The variants -(k)ko and -(f)fo often occur after vowels, but -(tʃ)tʃo does elsewhere (after vowels in geminated form and after consonants not geminated). However, more data is required to determine the exact phonological environment each of these prefers to occur.

Regarding plural formation, several variants are found; the following are the variants with examples:

a) Vocalic change: Vowel change has been recorded as plural marking as in:

-(w)a- to -ee-    man > meen- [person-women]

b) By adding the morpheme {-a} : man->mann-a; amboom-i>amboom-a;  
barra>barru-w-a.

We observe certain phonological processes during affixation: in man->mann-a, the root final consonant is geminated before adding the plural marking morpheme {-a}; in amboom-i>ambooma, the stem final /-i/ is deleted before adding the plural marker {-a}, and in barru-wa the final vowel of the singular noun barra, that is, /a/ is raised to /u/ and then the plural marker {-a} is added. It seems that diphthongs are not allowed in the language, and hence, a semivowel -w- is inserted to avoid vowel sequences.

c) Other plural markers observed include:

<b>(d)da-;</b>	<b>de; -</b>	<b>e;</b>	<b>he</b>
saa-da	wadʒɔʒu-de	doobb-e	beedda-he
kolli-dda			jema-he
duumma-dda			

It is not clear whether variations are allomorphic or areal. However, as they are found almost in all the four areas sampled, we consider them as allomorphic variations, though environment cannot be accounted by phonological conditions as all the variants, except {-e} which occurs after consonants, could be found after vowel endings.

**Areal variation:**

The morpheme {-wa} of Dara and Harbegona becomes {-ba} in Shebedino and Lokkaabaya. This means *w* is strengthened to *b*; or more probably and phonologically predictably *b* was the proto form and weakens to -*w* when it is not geminated and placed between vowels.

The plural marker {-de} of Dara is replaced with {-le} in Shebedino, Lokkaabaya and Harbegona. It seems that nouns whose singulative form ends with {-kko} tend to take {-he} as a plural form.

beedda- <i>kko</i>	beedda- <i>he</i>
star-SG	star-PL
'star'	'stars'
<i>jema-ko</i>	<i>jema-he</i>
mouse-SG	mouse-PL
'mouse'	'mice'

### Gender

In Sidaama, gender in nouns is marked by using suffixes and lexical changes; Table-6 below shows some examples of gender marking in nouns:

Masculine	Gloss	Feminine	Gloss
moti- <i>fffa</i>	'king'	moti- <i>tte</i>	'queen'
man- <i>ffu</i>	'a man'	man- <i>ffo-te</i>	'a woman'
beet- <i>u</i>	'boy'	beet- <i>o</i>	'girl'
haando	'ox'	lalo	'cow' (lexical)
gotf'ohu	'ram'	gere $\text{t}\text{f}\text{a}$	'ewe' (lexical)
faraa $\text{f}\text{f}\text{u}$	'horse'	faraa $\text{f}\text{f}\text{o- te}$	'mare'

Table-6: Gender marking in nouns

The masculine gender is basically marked by the suffix {-u}, but there are some variations as observed in *moti-fffa* [king-SG-M 'king'] and *haand-o* [ox-M 'ox'] in which the expected masculine suffix {-u} is lowered to middle back vowel {-o} due to the low back vowel /a/ in the stem noun. The feminine marker is {-te} as in *moti-tte*; combination of {-o} and {-te} is also used as in *man-ffo-te* [person-SG-F 'woman']. The irregular forms such as *gotf'ohu* 'ram' and *gere $\text{t}\text{f}\text{a}$*  'ewe' seem to have some element of gender marking as the former ends with '-u' and the latter has 'e' elements in the stem.



***Areal variation:***

It was found that *beetu* and *beeto* are used in Dara, but in Shebedino and Lokkaabaya, gender distinction is made with lexical modifiers *laba* 'male' and *n(m)eʔa* 'female'; hence, we find *lababeto* 'boy' and *neʔabeto* 'girl'. In this regard, *betto* is the default form rather than a feminine noun, and it often occurs with the short vowel /e/.

The derivational affixes of both nouns and verbs are largely shared, and hence we did not opt to present them here as the descriptions on derivation processes are presented in the works of Anbessa (2000) and Kawachi (2007).

**3. Conclusion**

There is a strong phonetic similarity (between 84% and 90%) among the varieties compared, implying low variation in phonetic features among the speakers of Sidaama. Lexical variation is slightly higher (similarity between 80% and 88%) compared to phonetic variation. The communication risk among Sidaama is low (12 %, the lowest risk and 20 %, the highest risk) implying an 80% intelligibility among Sidaama speakers. This in turn implies that Sidaama is a single language. The observed morphological variations are also limited and the native speakers know which form is found in which areas of Sidaama. As Sidaama is now taught in schools and is used in the regional media, the regional variations will tend to diminish with time and the language will move to a single standard language despite its contact with Gedeo, Oromo and Wolayta.

**Acknowledgment**

We would like to acknowledge the Research and Technology Transfer office of Hawassa University for sponsoring the research project under the thematic research Shared Values among Cushitic speakers in the South: Evidence from Biology, Culture and Linguistics.

### References

- Anbessa Teferra. 1987. Balliffa: Women's Speech among the Sidaama *JES*, vol XX, November 1987, PP.44-59
- Anbessa Teferra. 2000. *A Grammar of Sidaama*. Jerusalem: Hebrew University. Unpublished dissertation.
- Bayley, Robert and Lucas Ceil. 2007. *Sociolinguistic Variation: Theories, Methods, and Applications*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: Cambridge University Press.
- Beetana Hoteesso. 1983. *Sidaama People and Culture*. Addis Ababa: Bole Printing Press.
- Brøgger, Jan. 1986. *Belief and Experience among the Sidamo: A Case Study towards an Anthropology of Knowledge*. Oslo: Norwegian University Press.
- Casad, Eugene H. (1974). *Dialect Intelligibility Testing*. Summer institute of linguistics, Dallas Texas.
- Dukamo Duguna .2014. *Documentation of the Form and Function of Sidaama Demonstratives*. Addis Ababa University, Department of Linguistics (Unpublished MA thesis).
- Eshetu Tadesse W. 2010. *Sound Correspondences and Change: A Comparative Study of Burji, Hadiyya, Kambaata, Sidaama and Gedeo*. Addis Ababa University, Department of Linguistics and Philology (Unpublished Ma Thesis).
- Fekede Menuta. 2015. *Intergroup Communication among Gurage: A study on Intelligibility, Inter-lingual Comprehension and accommodation*. LAP Lambert academic publishing.
- Fekede Menuta. 2012. *Halabissa Grammar: Phonology, Morphology and Syntax*. Addis Ababa: Bole Printing Press.
- Gasparini, Armido. 1983. *Sidamo-English Dictionary*. Bologna: Editrice Missionaria Italiana.
- Holmes, Janet. 2013. *An Introduction to Sociolinguistics*. London and New York: Routledge: Taylor and Francis Group.
- Hudson, Grover. 2007. Highland East Cushitic Morphology. In Alan S. Kaye (ed.), *Morphologies of Asia and Africa*, vol. 1, pp. 529-545. Winona Lake, Indiana: Eisenbrauns.
- Hudson, Grover. 1976. Highland East Cushitic. In: Lionel M. Bender (ed.), *The Non-Semitic Languages of Ethiopia*. East Lansing, MI: African StudiesCenter, Michigan State University, 232-277.

- Kawachi Kazuhiro, 2007. *A Grammar of Sidaama (Sidamo), a Cushitic Language of Ethiopia*. ProQuest Information and Learning Company, Ann Arbor: USA.
- Schilling-Estes, Natalie. 2007. Sociolinguistic fieldwork. In Bayley, Robert and Lucas Ceil (eds), *Sociolinguistic Variation: Theories, Methods, and Applications*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: Cambridge University Press, 165-189.
- Tagliamonte, Sali A. 2007. Quantitative analysis. In Bayley, Robert and Lucas Ceil (eds), *Sociolinguistic Variation: Theories, Methods, and Applications*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: Cambridge University Press, 190-214.
- Tosco, Mauro. 2000. Cushitic Overview. *Journal of Ethiopian Studies* 33.2: 87-121.
- Tosco, Mauro. 2009. *Loan words in Gawwada, A Cushitic Language of Ethiopia: A Comparative Handbook*. Berlin: Walter de Gruyter.
- Treis, Yvonne. 2014. Number in Kambaata. In: Gerrit J. Dimmendaal & Anne Storch (eds.), *Number – Constructions and Semantics. Case Studies from Africa, Amazonia, India, and Oceania*. Amsterdam: John Benjamins, 111–134.
- Wedekind, Klaus. 1980. Sidaama, Gedeo, Burji: Phonological Differences and Likenesses. *Journal of Ethiopian Studies*, Vol. 14. 131-176. AA: Institute of Ethiopian Studies.
- Wedekind, Klaus. 1990. *Generating Narratives: Interrelations of Knowledge, Text Variants, and Cushitic Focus Strategies*. Trends in Linguistics: Studies and Monographs 52. Berlin: Mouton.