

Standardization of Oromo: Orthographic and Lexical Perspectives

Desalegn Leshyibelu Gelaglie

A Dissertation Submitted to the
Department of Linguistics and Philology

Presented in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Linguistics

Addis Ababa University

Addis Ababa, Ethiopia

June 2021

ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared by **Desalegn Leshyibelu Gelaglie** entitled: **Standardization of Oromo: Orthographic and Lexical Perspectives** and submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Tefera Negesse

Advisor

[Signature]

Signature

5-07-21

Date

Prof. Giorgio Banti

External Examiner

[Signature]

Signature

5-07-21

Date

GIRMA MENGISTU

Internal Examiner

[Signature]

Signature

05-7-21

Date

Chair of Department or Graduate Program Coordinator

Abstract

This study describes the orthographic and lexical standardization of Oromo. Qualitative research design was employed to describe the orthographic and lexical standardization of Oromo. Both primary and secondary sources of data were used. Focus group discussions and an interview were the primary sources of data; whereas the documents compiled by the standardization committee of Oromo and the primary and secondary school textbooks were the secondary sources of data. Even though there are a number of scholars who argue for and against the use of Roman based script, this study argues that Roman based script (*Qubee*) should be continued to write Oromo as a result of the linguistic, practical, acceptance and from the country's language policy point of views. Despite its occurrence in the various texts of Oromo, and its inclusion as an independent phoneme in the phonemic inventory of the language, the grapheme for the glottal stop is still not devised. Hence, this study, strongly recommends that it has to be represented with grapheme. Since the main reason for opting Roman based script is to mark the geminated realization of grapheme, this study argues that the sounds represented by the digraphs have to be marked when geminated. Though it requires experimental investigation, the graphematic representation of the ejectives <x> [tʰ], <q> [kʰ], <c> [tʃʰ] and <ph> [pʰ] may affect the maximum transfer of skills and may have a negative impact on the transfer of reading skills in English and Oromo as the graphemes with which these phonemes are represented are not the same in both languages. This pedagogical and transfer of reading skill challenges can be resolved via replacing the existing graphemes <x>, <q>, <c> and <ph> by <tʰ>, <kʰ>, <cʰ> and <pʰ> respectively. The geminated realization can better be marked via doubling only the first letter to be economical. The current alphabetic alignment of characters of the language is not systematic, particularly with regard to the order of digraphs in the alphabet. Hence, the study suggests the revisiting of the alphabetic order of the language. The finding also revealed that there are variations when writing Oromo ordinals, compounds, abbreviations, lexical and other word spacing related problems are the challenges of the standardization process of the language. The study believes that variation is due to the lack of codification and coordination among the stakeholders. Concerning the lexical elaboration strategy, both internal and external means are extensively applied in both the documents of the standardization committee of Oromo and in the textbooks to enrich the language. Semantic extension, derivation, compounding, blending, abbreviations, borrowing and loan translation are among the means which are used to elaborate

the language. Abbreviation and semantic extension are less productive in documents of the standardization committee of Oromo than in the textbooks. Semantic extension and borrowing and are more extensively used in the textbooks than in the documents of SCO. Meaning extension, derivation, compounding, blending, borrowing and loan translation are the most productive means of lexical elaboration in Oromo. All loanwords in Oromo are subject to modification. Regarding source language preferences, English is the main source of loanwords of Oromo as English is well developed to express scientific and technological concepts followed by Amharic, Arabic, Italian, Swahili and French. Greek loanwords are almost none in Oromo.

Acknowledgements

I came to this point in life; Glory to the Almighty God. Next, there are a number of people who have contributed to this dissertation and who deserve deepest gratitude.

My honest and deepest gratitude goes to my advisor Dr. Feda Negesse for his unreserved help from the beginning to the end of the dissertation. Dr Feda Negessa is not only my advisor, but also a mentor who exerted his maximum effort to deliver all the necessary reading material. Dr Feda, the energy, the initiation and motivation you extended to me helped me a lot in the journey to the dissertation. I have learnt a lot from you. Thank you so much Dr Feda. I would also like to thank Dr Shimelis Mazengia and Dr Samuel Handamo for comments and suggestions they forwarded during the upgrading stages of the paper. The valuable related materials they delivered as well as the comments and suggestions forwarded really helped to improve the paper. I am also grateful to Dr Awlachew Shumneka who has given me useful materials. I am very grateful to Addis Ababa University and my home institution, Dire Dawa University for sponsoring my PhD study. I am indebted to my home University staff for their help and appreciation on my journey to PhD study.

I would also like to thank Center for Oromo Culture for giving me all the necessary material for the study. I am very grateful to all members standardization committee of Oromo for their participation in the two phase focus group discussions. I am also grateful to educational experts who participated in the third focus group discussion.

I would also like to thank Dr. Derb Ado and department of linguistics for the opportunity I have had to participate in the summer course in Oslo University, Norway. I benefited a lot in acquiring research skill, financial and travel support. My heartfelt gratitude also goes to the staff members of Center for Multilingualism in Society across the Lifespan. I would also like to thank Professor Elizabeth Lanza, the center's director whose treatment motivated me to do better in my dissertation. My gratitude and appreciate also goes to Professor Pia Lane, Professor Francis Hult and Dr Cassie Smith-Christmas for the comments and encouragements they extended to strengthen my proposal of the project during summary course at Oslo University, Norway. They taught me in the Summer School of the University of Oslo.

The last, but not the least, my family also deserves gratitude and appreciation. My heartfelt gratitude also goes to Abeba Gebeyehu for the encouragement and unreserved help through the journey to PhD dissertation. I thank my mother Ehite Abire, my brother Yirgu Leshyibelu and his Wife Abebech Wegi for the encouragement extended to me all the years through.

Table Contents

Abstract	iii
Acknowledgements	v
List of Tables	xi
CHAPTER ONE	1
INTRODUCTION	1
1.1. Background	1
1.2 The language	5
1.3 Statement of the problem	8
1.4 Research questions	10
1.5 Objectives and rationale of the study	10
1.6 Significance of the study	11
1.7 Scope of the study	11
1.8 Methodology	12
1.8.1 Research design	12
1.8.2. Sampling	12
1.8.3. Data collection tools and sources	13
1.8.4 Data presentation and analysis	14
1.9. Ethical considerations	16
CHAPTER TWO	17
REVIEW OF RELATED LITERATURE	17
2.1 Conceptual framework	17
2.1.0 Language variation	17
2.1.1 Language planning	18

2.2 Review of studies related to standardization	48
2.2.1 Related studies on standardization of languages of other countries	48
2.2.2 Related studies on standardization of Ethiopian languages	51
2.2.3 Review of studies on standardization of Oromo	54
CHAPTER THREE	61
ORTHOGRAPHY DEVELOPMENT OF OROMO AND ITS STANDARDIZATION	61
3.0 Introduction	61
3.1. Oromo phonology	61
3.2. Syllable structure and vowel harmony of Oromo	65
3.3 Development of Oromo to a written language	66
3.3.1 Early attempts	66
3.3.2 Arguments over choice of script and the development of Roman-based script	69
3.4. Punctuation in Oromo	79
3.5 Numerals	80
3.6 Spelling convention	83
3.7 Evaluation of Oromo orthography	87
3.7.1 Maximum motivation for the learner and acceptance by the speech community	88
3.7.2. Maximum representation of speech	89
3.7.3. Maximum ease of learning	96
3.7.4 Maximum transfer of skill	98
3.7.5 Maximum ease of reproduction	99
3.8. Reform of Oromo orthography	100
3.9. Summary	105
CHAPTER FOUR	107

LEXICAL ELABORATION OF OROMO BY THE COMMITTEE.....	107
4.0 Introduction.....	107
4.1 Internal means of lexical elaboration	107
4.1.1. Semantic transfer method	107
4.1.2 Derivation	108
4.1.3. Compounding.....	117
4.1.4 Acronymization and Abbreviation.....	133
4.1.5 Blending in Oromo	135
4.2 External means of lexical elaboration.....	139
4.2.1 Borrowing	140
4.2.2 Loan translation	157
4.3 Summary.....	160
CHAPTER FIVE	163
THE ROLE OF EDUCATION TO THE (DE)STANDARDIZATION OF OROMO.....	163
5.1. Introduction.....	163
5.2 The role of education in the (de)standardization of Oromo.....	163
5.2.1 Lexical elaboration techniques	164
5.3. Comparison between lexical elaboration by the committee and textbook writers	252
5.4. Summary	254
CHAPTER SIX.....	256
SUMMARY, CONCLUSIONS AND RECOMMENDATION.....	256
6.0 Introduction.....	256
6.1 Summary.....	256
6.2 Conclusion	257

6.3 Recommendation	264
6.4. Potential areas of future research	266
APPENDICES	285
Appendix A. FGDs guiding and interview questions	285
Appendix. C. The recommended (by the present study) order of Oromo alphabet.....	292
Appendix D.Spelling variation in Oromo.....	294
Appendix E: Summary of loanwordsin Oromo	298
Appendix F Lexical variation due to dialectal variation (Sample).....	304
Appendix G: Transcripts of focus group discussions and interview	305
Appendix H: List of texts of the corpus.....	350
Appendix I: Summary of expertise of SCO who participated in FGD 1	352
Appendix J: Summary of expertise of SCO who participated in FGD 2.....	353
Appendix K: Summary of interviewee of head of center for culture of Oromia.....	353
AppendixL: Summary of expertise of education who participated in FGD 3	354
Appendix M: Short list of words codified by SCO.....	354

List of Tables

Tables	Pages
Table 1: The IPA equivalents of Oromo Qubee-----	15
Table 2: Language planning matrix-----	22
Table 3: Revised language planning model-----	23
Table 4. Consonants of Oromo-----	60
Table 5. Oromo vowels -----	61
Table 6. Consonant germination in Oromo-----	62
Table 7. Vowel Length in Oromo-----	62
Table 8. Oromo punctuations-----	78
Table 9. Oromo cardinals and their Arabic equivalent-----	79
Table 10. The most frequently used Oromo ordinals and their short forms-----	79
Table 11. Oromo ordinals and their short forms-----	80
Table 12: Recommended ordinal short forms of Oromo-----	81
Table 13. The current official orthography of Oromo-----	99
Table 14. The Recommended orthography of Oromo-----	100
Table 15. Inventory of compound types of Oromo based on word categories-----	122
Table 16. Inconsistently written Oromo compounds-----	123
Table 17: modified foreign loanwords in Oromo-----	136
Table 18: Adaptation of English vowels in Oromo-----	137
Table 19: Replacement of English diphthongs in Oromo-----	139
Table 20: Summary of loanwords in Amharic-----	243

List of Abbreviations, Acronyms and Symbols

1	First person singular
2	Second person singular
3	Third person singular
A	Adjective
ADP	Adposition
Amh.	Amharic
Bio.	Biology
CTBO	Culture and Tourism Bureau of Oromia
Ch.	Chemistry
COP	Copula
CPL	Completive (converb)
CSA	Central Statistical Agency
CS	Causative
CS ¹	Causative one
CS ²	Causative two
CS ³	Causative three
DAT	Dative
DF	Definiteness
EBO	Education Bureau of Oromia
E.C.	Ethiopian Calendar
Eng.	English
Env. Sc.	Environmental Science
EPRDF	Ethiopian People's Revolutionary Democratic Front
FDRE	Federal Democratic Republic of Ethiopia
FGD	Focus group discussion
G.C	Gregorian Calendar
GN	Genitive
Gr.	Grade

IDEO	Ideophone
IPV	Imperfective
LP	Language planning
M	Masculine
MD	Middle
MSG	Masculine singular
MoCT	Ministry of culture and tourism
MoE	Ministry of Education
N	Noun
NEG	Negation
NMNZ	Nominalizer
NOM	Nominative
Oro.	Oromo
P	Page
PL	Plural
PS	Passive
PV	Perfective
SCO	Standardization committee of Oromo
Soc.St	Social study
SG	Singular
VBZR	Verbalizer
Vs	Versus
//	Phonemic transcription or representation
[]	Phonetic transcription or representation
<	Results to
<>	Grapheme
→	Becomes, is realized as

CHAPTER ONE

INTRODUCTION

1.1. Background

The very notion of language planning means different things to different researchers in the field. This reflects the underlying ideological and theoretical fundamentals informing their conceptualization of language planning. For instance, as to Kaplan (2013:1), language planning is an activity, most visibly undertaken by the government, intended to promote systematic linguistic change in some of community of speakers. The author goes on explaining that the reasons for such change lie in a reticulated pattern of structures developed by the government and intended to maintain civil order and communication and to move the entire society in some direction deemed "good" or "useful" by the government.

To Jernudd (1971:3), language planning is a continuous work of giving solutions to the existing language problem in an organized way. Language planning is also an evaluation of linguistic change since language change is inevitable as the time passes (Haugen 1972:162). On the other hand, it is a deliberate action to affect a language and language users. For instance, Rubin & Jernudd (1971:2) claim language planning to be a deliberate language change characterized by evaluation of alternatives for solving language problems. Similarly, Cooper (1989:45) considers language planning as an intentional effort to influence the behavior of others concerning the acquisition, structure or functional allocation of their language codes. From Fishman's (1975:2) point of view, language planning is the authoritative allocation of resources for the attainment of language status and language corpus goals. This allocation of resources is either to make a language acquire new functions that are aspired to, or to make the language discharge its old functions more adequately.

Hence, we can infer some underlying motives of the above authors concerning language planning. Their arguments imply issues like a perceived language problem, various alternatives to choose from, deliberate language change, and goal orientation. These can be summarized into three major categories; namely linguistic factors, social factors, and political factors. The process of standardization is part of a larger phenomenon of language planning which Cooper argues LP

as a deliberate effort to influence the behavior of others with respect to the acquisition, structure and functional allocation of their codes (Cooper 1989:45). Language standardization, under the corpus planning activity, is among the basic elements of language planning activities.

Deumert (2004:2) treats language standardization as a process of variation reduction that includes deliberate intervention by regulating authorities (such as language societies and academies, individual dictionary and grammar writers and also government institutions; i.e. the imposition of uniformity through authoritative acts). Standardization is concerned with linguistic forms such as corpus planning that consists selection and codification as well as the social and communicative functions of language (status planning, i.e. implementation and elaboration).

Standardization, from Deumert's (2004:2) perspective, could be mono-centric or polycentric. Mono-centric standardization is a kind of language standardization in which one variety of a language becomes widely accepted throughout a speech community as a supra-dialectal norm (Ferguson 1996:43; Deumert 2004:2). On the other hand, a polycentric standardization is a complex recombination of features from various varieties. The main purpose of standardizing a language is to construct belief and practice that there is a common linguistic ideal to which language users within a determined community should aspire. This is due to the concern that all members of a society should have universal access to a common variety of a language. Language standardization is also for the need of uniformity and emphasizes and respect for centralized language variety while backgrounding the values of linguistic diversity.

Authors often discuss the language planning and policy of Ethiopia from the three era point of view; i.e., the imperial system, the socialist system and the federal system. Concerning language policy before the Derg regime, the language policy difference observed between Haile Sellasie and his predecessors was that the *de facto* national status of Amharic was confirmed by *de jure* as a law in the Haile Sellasie's constitution (Samuel 2017:13). Amharic was also used as a medium of instruction in the primary education throughout the country during the imperial system (Tesfaye & Taylor 1976:373; Zelealem 2012:9). This trend had also continued till 1993.

On the other hand, when we compare language policy from regime to regime, it is clear that Ethiopia has shown a significant change in terms of developing and implementing language policy from the reign of Tewodros to date. As to Getachew & Derib (2006:58), the covert language policies that were implemented by Tewodros II, Minilik II, Haile Sellasie I and the Derg all, more or less, promoted the use of one language, Amharic. The authors also further claim that the language policy that the Derg developed seemed to give a hope for the development of many Ethiopian languages, but ended up in vain as the implementation was just a continuation of the use of one language. At the beginning of the regime, the Derg stated that all nations, nationalities and ethnicities to enjoy the freedom of maintaining, using and developing their language together with many political and economic benefits (Getachew & Derib 2006:47). Accordingly, during the National Literacy Campaign took off in 1975, fifteen local languages; namely Amharic, Oromo, Tigrinya, Tigre, Wolaitta, Sidaama, Haddiya, Kambaata, Afar, Saho, Gedeo, Somali, Kafinono, Silt'e, and Kunama were used for the literacy purpose which lasts till 1990 (Zealelem 2012:24; Tilahun 1997:507). Even the Derg regime was better than its predecessor in at least introducing more languages during the literacy campaign, the policy still favored the majoritarian or dominant interests at the expense of minority and non-dominant interests which violates the critical approach of language policy (Zealelem 2012:24).

During the Ethiopian People's Revolutionary Democratic Front (EPRDF), language policy towards the implementation of a multilingual language policy came into practice since 1994. More than 30 languages are now being used as a medium of instruction in primary education and some languages like Oromo, Amharic and others are being taught as subjects from primary to tertiary education (Vujich 2013:5). Books, articles, newspapers, radio and television broadcasts are available in a lot of languages as well.

Language Planning and policy related issues are also included in the constitution of Federal Democratic Republic of Ethiopia and in other policy documents ratified by the Ministry of education and the Ministry of Youth, Sport and Culture to make the ground fertile for the implementation of multilingualism and language development into practice. Hence, discussing language planning and policy related articles of the documents are very important for the present

language standardization study. Article 5 of the 1995 constitution (FEDRE 1995) reads as follows.

1. All Ethiopian languages shall enjoy equal state of recognition
2. Amharic shall be the working language of the federal government; and
3. Members of the federation may determine their respective languages

Following the constitutional right, various regional governments have decided which language(s) to use as official working language, medium of instruction in primary schools and in the media. For instance, regions like Oromia, Tigray and Somali which have relatively homogeneous regions have opted for the dominant language of the area for the purpose of regional working language. To the contrary, linguistically diverse regions such as Southern Nations, Nationalities and Peoples Region and Benishangul Gumuz opted for Amharic as the language of regional communication (Cohen 2007:66; Samuel 2017:15).

The Ministry of Youth, Sports and Culture also stated language related policy (MoCT 1995). Section 1.4 of the document states that all the languages and dialects of the country shall be identified, inventoried and classified in accordance with their geographical distribution and standards (Ministry of Youth, Sports and Culture 1995). This section of the document states corpus planning activities of a language. Section 6.1 of the document also states that a language distribution map of the country shall be prepared. This section of the document is all about the necessity of the study of the distribution of a language (and its variety) in the country. This could also be a base for making language standardization related activities. There are also other language standardization related statements in the policy document. For instance, section 6.2-6.4 of the document (Ministry of Youth, Sports and Culture 1995) states that alphabets shall be developed to those languages that do not have scripts and put them into service taking into account the psychological attitude of the speakers. This document clearly shows that graphization is important to reduce unwritten languages to a written language so as to serve its society. Moreover, section 6.3 of the document states the development of alphabet in itself is not enough; up-grading measures shall be taken to solve the problems of written languages in accordance with the principles and methods of the science of orthography. From this section of

the document, we can infer that orthography of a language has to be reconsidered to check whether it is designed in accordance with scientific principles of orthography.

The standardization of orthography alone may not normalize a language. As a result, Ministry of Youth, Sports and Culture (1995), in its document under section 6.4 states that there has to be also an enabling environment in order to prepare dictionaries, encyclopedia and grammar texts in the languages of the various nations and nationalities of Ethiopia. The Ministry of Youth, Sports and Culture gave due attention (at least in written) to scientific and technological terminologies. Standardization of scientific and technological terminologies can help in promoting the capacity of the languages of the country (see section 6.6 of the Ministry of Youth, Sports and Culture (1995)). There are, however, shortcomings in putting all those normalization activities so as to solve the problems of written languages of the country and the case of Oromo is not exceptional.

The recently ratified language policy of Ethiopia clearly stated in its fundamental principle section of the policy that the working languages of the federal government shall be Amharic, Afan Oromo, Tigrigna, Somaligna, and Afarigna (MoCT 2020:14). It is also stated in the policy that other Ethiopian languages have a potential to become, in due time, the working languages of the federal government based on the availability of the capacity of the country. Further more, Ethiopia also acknowledged all languages and their writing systems as national heritages and assets (MoCT 2020:7). But, the standardization process of the languages of the country still needs a plan and an action.

1.2 The language

Oromo is a Lowland East Cushitic language of the Afro-asiatic phylum. The people call themselves Oromo and their language Afaan Oromoo. Various scholars, however, often use different names for the language: Oromiffa, Oromic, Oromo, Afan Oromo and Oromo language (Tekabe 2010:1). This research opted Oromo to refer to the language and the people as it is used by many authors in various literature.

Regarding the linguistic groups in Ethiopia, Oromo takes the lion's share. Oromo covers a large area that stretches from the Sudan border in the West, through Addis Ababa, and beyond Harar in the East. It also extends up to Northern Kenya in the South and East of the rift valley, and to

Wollo in the North (Gragg 1982: xiii). Concerning the population size of Oromo speakers, various scholars state various figures. As stated by Mahdi (1995: xi), the number of the Oromo people in Ethiopia is close to 30 million. The 2007 Population and Housing Census of Ethiopia has reported that the population size of Oromo is 24,930,424 (33.8%) of total population of Ethiopia. Mohammed (1994:7), argues that the Oromo people constitute at least fifty percent of the Ethiopian population.

Oromo is a language of administration in the region. It also serves as a language of media and education in the Regional State of Oromia. Various materials such as books, newspapers, magazines and so forth have been published in different scripts; though the language had remained without an officially accepted system of writing until 1991. Despite lack of a convincing reason for their choice of different scripts for Oromo writing, the influence of the early Oromo writers on the immediate and later generations has been in evidence by the fact that Oromo has continued to be written in various scripts until the beginning of 1990's. The Roman based script, Arabic script and Ge'ez script were among the scripts used by various authors (Teferi 2015:14) until the stated years; but 'Qubee', which is a Roman based script, has been officially accepted since then.

There is no consensus over the classification of Oromo dialects as various authors classified the dialects in different ways. For instance, Bender, Mulugeta & Stinson (1976) classified the dialect into seven; Macha, Tulama, Wollo, Rayya, Arsi, Guji, and Borana. However, Wako (1981) classified them into five groups; Southern (Arsi, Guji and Borana), Central (Karayu, Selale), Mecha (Jimma, Wollega and Ilubabor), Eastern (Harar and Bale) and Northern (Raya and Wollo).

Gragg (1982:2), on the other hand, classified the dialects of Oromo Western (Wallagga, Iluu Abbaa Bor, Jimma), Central (Showa), Eastern (Hararge) and Southern (Arsii-Baale, Gujii and Boorana), the Baate and Raayyaa of Wollo and Tigray. As to Kebede (1997:3), the dialects are North Western (Tulma, Mecha), Eastern (Harar, Arsi-Bale, Wallo, Rayya), Central (Arsi-Zeway, Guji, Borana, Munyo, Orma) and Southern (Waata). In subsequent study Kebede (2005:2), again resorted the dialects in to five as follows.

1. Waata, Orma, Borana of Ethiopia and Kenya, and Arsi,
2. Tuulama and Macca,
3. Raya,
4. Hararge
5. Baate

Kebede (2007:5), again reclassified the dialects into 14; Wollo, Raya, Tulema, Mecha, Arsi, Hararge, Guji of Borana in southern Ethiopia and northern Kenyan, Orma, Gabra, Ajuran, Sakuye, Garreh, Munyo, and Waata. Banti (2008:5) also classified the dialects of Oromo into six clusters; Northern Oromo (Baate and Raayyaa), Western Oromo (Macca), Highland Showan Oromo (Tuulama), Eastern Oromo (Hararge), Central Oromo (Gujii and Arsii), and Southern Oromo (Boorana).The dialects of Oromo are reclassified into ten by Kebede (2009:1) for the fourth time using the genetic classification method as Western, Eastern, Central, South-East-North, Waata, North-East, East, North, Wolo, Ray. Feda (2015) classified the dialects into six (west (Wollega, Jimma and Illuababor), central (Showa), northern (Rayya, Wollo), southern (Borena and Guji), southeast (Arsi and Bale) and eastern (Harar)) by calculating the Levenshtien distance among the dialects on the basis of the distance values.

The dialects of the language are classified mostly on the basis of geographical distribution as the Oromo people are scattered over a large area of Ethiopia and some parts of Kenya and Somalia as well (Stroomer 1995:1).Stroomer (1995:2-3) adds Boranas spoken in Kenya, the eastern part, Sakuye, southern part, Garreor Gurre, north-eastern part and Munyo, close to the Orma dialect (Ali & Zaborski 1990: ix). Avanzati & Gutman (2013:14) discussed dialects of Oromo spoken on the border of Ethiopia and the neighboring countries like Somalia and Kenya. The Gabra dialect is spoken in a cross-border area between Ethiopia and Kenya, Orma and Waataa dialects are spoken in Kenya.

As can be observed from the above classification of the dialects of Oromo, the existence of wide ranges of variation in Oromo necessitates standardization of Oromo (Teferi 2015:274). Following this necessity, the Oromia Culture and Tourism Bureau has established a language

standardization committee. This committee has been working on the standardization of the language for more than two decades.

1.3 Statement of the problem

Several tasks have to be performed in the various fields of a language so as to realize standardization of a specific language. The creation of a new script to unwritten language or changing the existing orthography and alphabet for a written language is an important beginning task of standardization of a language. The standardization of the designed script, coining new terms, reforming spelling, creating new forms, the modification of old ones, the selection standard forms from the alternative forms (corpus planning in general) are very important for the attainment of a standard language (Cooper 1989:125; Ndimande-Hlongwa 2010:207). Codification is also another essential element of language standardization. In this regard, Trudgill (1999:117) treats codification as a process by which a dialect of a language can be recognized and confirmed officially by language planners, or policy makers to be used across the other dialects of the language. The codified forms are usually included in the guide manuals, dictionaries and grammar books in order to be used widely and spread among speakers of different dialects. The selection of base dialect, orthography, lexicon etc., can be decided by a rule (Crystal 1997:67). The criteria for setting norms and conventions for orthography of a language in polycentric approach, according to Tauli (1974:60) and Haugen (1966:177), are the number of speakers, frequency of occurrences, originality, economy, semantic transparency, acceptance, written documents, productivity, the use in the mass media and current status of features of the language. Hence, investigating the standardization of the language from language standardization point of view is important.

Coming back to the present study's concern, Oromo is relatively one of the most studied languages; particularly with respect to its description; but a lot remains concerning language standardization (Teferi 2015:269). Mekonnen (2002) examined lexical standardization in Oromo in terms of the four categories of the process of standardization, namely; selection, codification, the elaboration and implementation and proposed all varieties as a base for the standardization of Oromo. The study describes the standardization process based on Haugen's (1966; 1969) conceptual model of language planning, i.e., selection, codification, elaboration and

implementation. Teshome (2014) examined lexical use of Oromo in the government mass media. Baye (1994) examines lexical development in Oromo. None of them examined the current orthographic and lexical standardization of the language (see section (2.2.3)).

The present study, however, chiefly focuses on the orthographic and lexical standardization of Oromo. To the best of my knowledge, there is no detailed and up to date study that shows the current standardization status of the language. Despite the existence of some orthographic related issues to be reconsidered, the orthography of the language has not yet been reformed for the last 27 years. Oromo is expanding its vocabulary to express scientific and technological concepts. Standardization of the language by the standardization committee of Oromo is still an on going process. There is no study regarding the current elaboration techniques being applied by the standardization committee and the textbook writers. There is also a research gap regarding the current challenges of the standardization of the language. Hence, investigation of the orthographic and lexical standardization of Oromo is important. Although orthography is usually subject to demand for improvement for the sake of simplicity, learnability, linguistic economy and for all other principles of orthography, changing an established orthography is extremely difficult, since conservative elements who are in favor of maintaining the status quo resist any change (Coulmas 2013:108ff; Sebba 2009:133; Shimelis 2016:215). In this regard, to the best of my reading, I have not come across any orthographic reform of Oromo by the standardization committee of the language since the Roman based orthography has been officially used to write the language.

The identification of codified rules, conventions and enrichment techniques of Oromo by the standardization committee of the language are still untouched area of study. On the other hand, education sector plays a paramount role more than any other institutions as it is a site where a strict regulation of forms is applied and taught. In this regard, the language policy of Ethiopia seems to support this concept. The policy, under the section of the writing systems of the languages of the country states that a standard writing system shall be designed for the formal function of any language of a nation, nationality or people (MoCT 2020:23). The policy also clearly states that government institutions, educational institutions and the mass media shall have special responsibility to maintain the standards of Ethiopia's writing systems that are currently in

use (MoCT 2020:23). In this regard, this study is going to identify whether the education sector (textbook writers) and the standardization committee of Oromo are working hand in hand or not. Hence, so far, there is no evidence regarding the current standardization status of Oromo. To the best of my knowledge, there is, again, no sufficient study that fully addressed the current practices of lexical elaboration of Oromo in the educational domain (particularly in the primary and secondary textbooks prepared in Oromo) and in the materials compiled by the standardization committee of Oromo. The existing challenges of standardization of Oromo are also still the area that needs investigation.

Hence, it would be important to scientifically investigate the linguistic techniques for lexical elaboration in Oromo. It is also important identifying the orthographic and lexical standardization of the language. Therefore; the present study is chiefly concerned with investigation of orthographic and lexical standardization of Oromo.

1.4 Research questions

Hence, the study is aimed at answering the following research questions:

1. How can the orthography of Oromo be evaluated against the principles of orthography standardization?
2. How is Oromo being codified by the standardization committee of the language?
3. What are the linguistic techniques applied to elaborate the language?
4. What are the challenges of standardization of Oromo?

1.5 Objectives and rationale of the study

The general objective of this study is to describe the orthography and lexical standardization of Oromo. Specifically, the study intends to:

1. Evaluate the orthography of Oromo against the principles of orthography standardization
2. Examine the codification process of Oromo.
3. Identify the linguistic techniques applied to elaborate the language.
4. Identify the challenges of standardization of Oromo.

The main reason why I chose to work on the standardization of Oromo is due to the fact that there are only few works on the standardization related topic, particularly on Oromo and due to the interest that I have to contribute to the existing knowledge. The rationale of preferring Oromo is due to two reasons. The first reason is that Oromo is one of the languages of Ethiopia that shares a border with other countries and almost with all the borders of the states of the country. Since geographical broadness entails language variation, the language has a potential to have challenges of standardization which is my field of interest. The second reason, though subsidiary, I am a native speaker of Oromo (central variety) which makes me advantageous and that may add flavor to the study.

1.6 Significance of the study

The study is believed to offer necessary information to the concerned bodies, particularly to ‘*KoreeWaaltina Afaan Oromo*’ (Oromo Standardization Committee) about the standardization situations of the language. The education sector personnels and textbook writers can find the result of this study useful as the area is where regular language usage is strictly obeyed. It can be a good input for future researchers, especially for those who are interested to work on language standardization issues. The language planners working in academy and individual levels may find the result of this study useful. This particular work may also be helpful for those who are involved in media, orthographic development, lexicography and translation.

1.7 Scope of the study

The study is limited to the description and analysis of orthographic and lexical standardization of Oromo. It is also limited to the description of the practical challenges of orthographic and lexical standardization of Oromo. The source of data of the study is also restricted to the materials published by the standardization committee of Oromo and from the current primary and secondary school textbooks published in Oromo.

1.8 Methodology

This part of the study discusses the methodology of the research. The research design, sampling techniques, data collection tools and sources of data, the data presentation and analysis techniques are discussed in the following consecutive sub topics.

1.8.1 Research design

Both qualitative and, in few cases, quantitative research design were employed to describe the orthographic and lexical standardization of Oromo in the documents compiled by the standardization committee of Oromo and in the primary and secondary school textbooks. A frequency count and a percentage were used to quantify loanwords and loanword source languages as well as to show spelling variation of lexical item in the Oromo.

1.8.2. Sampling

Since the main concern of this study is on the orthographic and lexical standardization of Oromo, the major sources of data were written materials (text corpus). The standardization committee of Oromo (an authoritative and a legitimate governmental body) and education sector were selected as the main domains since these sites are where the task of codification (normalization) process takes place and where standard forms are strictly applied respectively. Hence, the materials compiled by the standardization committee of Oromo were selected and became secondary sources of data. There are 11 *Wiirtuu* books and one grammar book which were compiled by the committee and that are related to the standardization of the language. All of them were selected and included in the corpus of texts and were used as a source of data. From the education domain, all the current primary and secondary school textbooks which were published in Oromo were included in the corpus and were sources of data. Accordingly, there are 52 textbooks of primary and secondary schools which include all subjects (see appendix H). All these textbooks were parts of the corpus. The materials published by the standardization committee of Oromo and primary and secondary textbooks are purposely sampled as they are the core areas where standardization and implementation processes take place respectively. Textbooks and schools in general, are areas where a strict standard (uniformity) of a language is followed and practiced. Finally, based on purposive sampling strategy, the head of center for Oromia culture was

selected to participate in the key research participant interview session to collect data related to standardization of the language.

Apart from the text corpus, three focus group discussions (FGDs) were also conducted; two of which were on the site of the center for Oromo culture and one FGD at Oromia education bureau. All (there are 12 personnel) members of the standardization committee of Oromo were included in the focus group discussions in two sessions. Accordingly, six members in the first FGD and the rest six members of the standardization committee of Oromo were included in the second session. The third, focus group discussion also includes six participants who are experts working in the education bureau of Oromia. These research participants were selected using systematic random sampling techniques from the list of 24 education experts. Accordingly, every 4th expert was selected. The selected research participants were from different fields of study; namely from language (Oromo), civics and ethical education, educational leadership and management, chemistry and mathematics. Most of these participants participated during textbook preparation (see appendix G, III).

1.8.3. Data collection tools and sources

Selecting the appropriate data collection tools is crucial in order to obtain the intended data and thereby to attain the objective of the study. Accordingly document analysis, focus group discussions and interview were used to collect the data. Regarding sources of data both primary and secondary sources of data were used under the present study. The primary data are drawn via conducting focus group discussions and interview. Secondary data emanate from the written documents of the standardization committee of Oromo and from a corpus of textbooks that includes all (52 textbooks from grade 1 to grade 12) primary and secondary textbooks published in Oromo. The primary sources of data emanated from the three focus group discussions and an interview. Here, since the majority of the data were drawn from the secondary source of data (from materials compiled by the standardization committee of Oromo and primary and secondary textbook corpus), the primary sources of data i.e., focus group discussions and interview mainly supplemented the secondary sources of data. In a few cases, I have employed my own native

speaker introspection and the consultation of the native speakers of varieties of Oromo whenever I found it necessary.

1.8.4 Data presentation and analysis

The present study chiefly applies a qualitative method of data analysis. Orthographic and lexical data that were extracted from the corpus were presented orthographically and arranged in theme by theme. The data that were collected through content analysis of the selected documents were analyzed by grouping into thematic areas like orthographic standardization and lexical elaboration and variation. Those data which emanate from the education domain were presented and analyzed separately to compare the orthographic and lexical standardization in the materials published by the standardization committee of Oromo and in the textbooks. The analysis was also supported with sufficient illustrative examples for each thematic area. Orthographic transcription was also employed to present qualitative data that were gathered through interviews and focus group discussions. Primary and secondary data were presented and analyzed together.

In the present study, data are presented using Oromo orthography except for /ʒ/ since there is no grapheme representation for /ʒ/. Geminated consonants and long vowels are indicated by doubling the symbols as it is an orthographic rule of the language. As the study focuses on standardization of the language, the orthographic transcription has been used and IPA equivalents of Oromo orthography are given in the following table. The phoneme, grapheme and alphabetic order below are based on Qajeelcha (1998:14) except for the glottal (‘).

Table 1. The IPA equivalents of Oromo Qubee

Grapheme (Oromo Qubee)	IPAequivalents	Grapheme (Qubee)	IPAequivalents
a	/a/	r	/r/
b	/b/	s	/s/
c	/tʃ/	t	/t/
d	/d/	u	/u/
e	/e/	v	/v/
f	/f/	w	/w/
g	/g/	x	/tʰ/
h	/h/	y	/j/
i	/i/	z	/z/
j	/dʒ/	ch	/tʃ/
k	/k/	dh	/dʰ/
l	/l/	ny	/ɲ/
m	/m/	ph	/pʰ/
n	/n/	sh	/ʃ/
o	/o/	'	/ʔ/
p	/p/		
q	/kʰ/		

1.9. Ethical considerations

To Stern & Elliott (1997:4), ethics that must be practiced in doing any research is that the researcher must state the necessary information to the participants before asking them to participate in the study. The researcher first secured the required permission from authorities in the study areas to conduct the study after presenting a letter of cooperation written by the Department of Linguistics and Philology. Prior to collecting data from each participant, the researcher explained them the purpose of the study to obtain their consent. The researcher asked the participants to give their consent to participate in the research and the answer was positive. The researcher used a voice recorder during data collection after securing the consent of the participants and confidentiality of data, names and other personal information are secured. These ethical issues are considered in the present study.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

The review of related literature consist two parts; conceptual framework and review of related studies. The first part discusses the conceptual framework of the study that is related to language planning in general and language standardization in particular. The second part of the literature deals with the studies that are related to the description of language standardizations.

2.1 Conceptual framework

2.1.0 Language variation

Language variation occurs due to language contact or due to the way people use their language or because of the contact with other languages and cultures (Rosa 2017:36). Language variation could be synchronic and/or diachronic or historical. Synchronic variation is a variation that may arise due to geographical variation or social variation (Penny 2012:1). Geographically, speakers of the ‘same’ language who live in different parts of a continuous territory do not speak in the same way (Penny 2012:1).As a result of existence of distances between geographical area successive small differences accumulates as one crosses an area. This accumulation of differences may result in total mutual incomprehensibility between the speeches belonging to distant parts of the territory (Penny 2012:1). Language variation could also be due to social variation that is triggered by factors that include age, sex, race, class background, education, occupation, and income. Generally, all aspects of language (sounds, phonemes, morphemes, syntactic structures, lexemes, meanings, etc.) are subject to variation according to geographical and social parameters (Rosa 2017:41). Wolfram (2006:333) summarizes the issue of language variation as the structure to be the heart of language and variation to define its soul.

A language comprises a number of diverse varieties that includes dialects, style, register as well as the standard variety. Herbert (1992:3) states that each language label presents a range of varieties. As per the author, it is necessary to distinguish the standard variety from other types and to look at the role of this standard variety within a language. Crystal (1984:337) and Romaine (1984:84) discusses the concept of *style*. According to them, *style* can be defined as different

ways of speaking adopted by speakers in different situations. For instance, the situation or circumstance in which the speaker finds himself or herself can govern his or her way of speaking (Wardhaugh 2002:48). The variety is then described as situational. Romaine (1994:84) points out that the most common influence is the degree of formality. This implies that the choice of language to be used by the speaker is determined by the situation. The speaker will use formal language in formal situations, and informal language in informal situations. For instance, the way a principal at a school addresses or speaks to the teachers in a staff meeting will differ from the way he/she would speak to them at home or in an informal conversation. The situation in which the language is used allows the use of certain words and prohibits the use of others. As the situation becomes more formal, the speaker's behavior and the use of words also become formal. Variation could also refer to a specific feature of a language which shows a particular variation in a community and which is used as a tag for classifying a speaker's speech. For example, in eastern Oromo the realization of /k/ is just /x/. Such variation also tends to happen in a written language too. The variation could be orthographic, phonological, word spacing, lexical and semantic. There are such variations in the language. A single concept seen to be represented with more than three forms as in, for instance, *Wixata*, *Dafino*, *Hojjaa duraa* to refer to 'Monday' (see section 5.2.1.2). Variation can be minimized through a process of standardization of a language by the authoritative body.

2.1.1 Language planning

The term LP was first used by Haugen (1959:8), who defined it as the activity of preparing a normative orthography, grammar and dictionary for the guidance of writers and speakers in a non-homogeneous speech community. Haugen (1983) emphasizes selection, codification, elaboration and implementation as main activities in LP. Various scholars discuss language planning from different points of view. Language planning (hence forth LP) is usually defined as a deliberate effort to influence the function, structure or acquisition of a language or language variety within a speech communities or any political attempt to change the status of a language in some way or develop new ways of using it (Cooper 1989:45). Rubin & Jernudd (1971:3) argue that LP is a deliberate intervention in the process of language change. They further state that LP is manipulation of the linguistic resources of a society to achieve educational, political, and economic ends of multilingual states. As to Deumert (2004:45), it is a systematic attempt to

solve communication problems of a community by studying its various languages and dialects and developing an official language policy regulating their selection and use. LP is carried out by those who exercise political power. Language planners and policy makers often face the question of accommodating the needs of linguistically diverse communities.

Reagan (2002:419) states that LP involves public decisions about language, its use, status and development that have overwhelming social, economic, educational and political significance for the society and individuals. Reagan (2002:420) says that LP is a conscious and future-oriented activity involving choices on decision-making processes. It can serve as a tool for empowering groups and individuals, for creating and strengthening national bonds and ties, and for maximizing educational and economic development. Reagan (2002:420) also mentions that it is used to defend domination, social-class discrimination, and social and education inequity. Adebija (2004:192) states that language policy and planning decisions that arise in response to the socio-political needs of the speakers. Its decisions may be required, for example, when a number of linguistic groups compete for access to day-to-day life, or when a particular linguistic minority does not get such opportunities. LP decisions typically attempt to meet these needs by reducing linguistic diversity, for instance, by declaring a single language a national language in a multilingual country, or by selecting a single variety of a language as “standard” to promote linguistic unity in a country where divergent dialects exist (Tollefson 2002:3 & Tiglu 2015:41).

Tollefson (1991:22-42) mentions that LP efforts are inevitably ideological in nature. It is done through the cooperative efforts of political, educational and linguistic authorities. Language planners are involved in the selection of official language, development of writing systems, preparation of dictionaries, grammars, and textbooks, as well as in the promotion of literacy and standardization, modernization, and terminological enrichment of majority and minority languages.

LP also entails the assessment and alteration of the practices and products of government, private business, and the media (Cooper 1989:184). LP involves coining words and thinking up spelling reforms. It covers a mixture of approaches, including terminology and lexicography, terminology management, translation and translation management. Bamgbose (1991:109) claims that

language planners and policy makers are faced with major challenges of how to formulate and implement language policies in a way that minimizes the incidence of exclusion. Mutasa (2004:24) presents LP as a concept that identifies language problems, describes them and influences change with the aim of solving the problems. Besides, LP in a multilingual context aims to improve access for children to formal education by facilitating the implementation of mother tongue education particularly at pre-school and primary school levels. It enhances cultural diversity in a country by fostering the arts in the different languages. It also prevents ethnic and political dissatisfaction on the part of the people by creating an environment based on the principle of language equality, especially in official sectors of society such as legislation, justice, public administration and education, bearing in mind the different developmental stages of co-existing languages (Mutasa 2004:65).

Although language planners believe that there is no theory of LP, they make tremendous use of models of LP for problem identification, issue definition, agenda setting, policy formulation, implementation and evaluation. For instance, Rubin (1971:218) developed LP model which has four steps (i) fact-finding, (ii) policy formulation, (iii) implementation, and (iv) evaluation. Language planners who work within this framework present it differently. However, there seems to be agreement on its key facts.

To Rubin (1971:218), the first stage is fact finding as LP should be grounded by facts which focus on identifying problems and establishing goals. Hence, the planner must determine the rationale, the existing social, cultural, political and economic dimensions. It is imperative to do a thorough fact-finding process in LP, particularly in linguistically complex countries like Ethiopia. LP, especially when empowering indigenous languages, can initiate social change which might be resisted because of power, the financial cost, ethnic differences, and individual dialectal preferences (Rubin 1971:218). According to Mutasa (2004:26), fact-finding should establish three types of basic data; attitudinal, demographic and situational to determine the success or failure of LP. It, therefore, is an attempt to predict what occurs if certain aspects of language use are modified. This planning stage includes brainstorming about how a linguistic problem can be approached, and how the selected solution is preeminently implemented.

The second stage, policy formulation, is the actual planning whereby an action plan will be formulated based on an understanding of the constraints (Rubin 1971:219). At this stage, the language planner decides on goals and strategies, and predicts outcomes based on the information gathered through fact-finding. They also consider the possible influence of the government in their plans. The language planner may modify the initial plan to adopt it to the actual situation. One challenge the planner may face in multilingual countries is to find a balance between non-linguistic and linguistic goals.

The third stage in Rubin's (1971) model of LP is the implementation phase. Mutasa (2004:26-27) acknowledges that this is the most challenging stage, which calls for the cooperation of all stakeholders. Language planners have the task to gain the consent of the majority of the population, and the government. The challenge here is that what may be acceptable to the population may not be acceptable to the government or vice versa. The final stage in Rubin's (1971) model is evaluation. At this stage, the planner assesses if the actual outcomes match the predicted outcomes (Rubin 1971:220).

Batibo (2005:118) advances two aspects of LP, ideological planning and technical planning. Ideological planning could be equated to status planning since it involves policy decisions on what language should be used for government, business, public administration, education, mass media, diplomacy, international relations, legal matters and commerce. According to Bamgbose (1991:111), technical planning involves norm planning, codification and capacity planning, which requires the intervention of language experts; especially on the codification and capacity planning.

Cooper (1989:98) considers LP in terms of eight components: what actors attempt to influence which behaviors of which people for what ends and under what conditions by which means through what decision-making process and with what effect. While many models of LP focus on the nation or state as principle actors, LP is also undertaken by other agents at the sub-national or -state levels or the supra-national or -state levels. Cooper (1989:115) states the behaviors which LP seeks to influence can focus on the linguistic features of a language or on its social qualities.

Finally, Cooper (1989:31) argues that any LP needs to consider the decision-making processes since the way in which measures are decided upon influences its effect. The process of LP is an ideological and political activity since it is policy driven (Ball 2005:120). Top-down language planning is a type of language planning in which activities are done by only experts without involving other stakeholders like speakers. Top-down imposed LP is seldom effective if it does not enjoy support among those for whom it is intended (Spolsky 2004:222; Wright 2004:169). On the other hand, bottom-up language planning is a kind of LP in which all stakeholders participate while planning the language and then what is planned will be approved by experts. However, Kaplan & Baldauf (1997:14) note that LP is everybody's business, and without the help of the communities, language ecology will not develop and no amount of planning is likely to bring sustained language change. It is the right of ethnic groups to maintain their cultures and beliefs (Ricento 2006:17). The above concept denotes that LP stakeholders can play a great role to implement language related decisions.

On the other hand, Haugen (1966a:933) states that there are four activities in language planning: selection of norm, codification of form, elaboration of a function, and acceptance by the community. For Haugen (1966b:17), *form* is the linguistic structure in its ramification and *function* is the variety of the uses to which the structure is put. Haugen states that *selection* and *acceptance* as societal aspects of language development, whereas *codification* and *elaboration* as linguistic aspects. Table 1 shows the language planning matrix of Haugen (1966a:933).

Table 2: Language planning matrix

	Form	Function
Society	Selection	Acceptance
Language	Codification	Elaboration

(Source: Haugen, 1966a:933)

Language planning model of Haugen (1966a:931) focuses on the standardization and functional elaboration of languages within a speech community. The model focuses to minimize linguistic variation via codification and functional extension of the codified variety through elaboration. The agreement of the speech community on the selection of some kind of a model from which

the norm can be derived is a crucial point for the minimization of the variation and elaboration of functions (Haugen 1966a:931-932).

Haugen (1983) substituted the term acceptance by implementation in his revised model due to the fact that the acceptance or the rejection of the speech community is one of the critical issues to be considered during language planning. As shown in Table 2, in the new model, the societal dimension is specified as *status planning*, but the language dimension as *corpus planning*:

Table 3: Revised language planning model

	Form (Policy planning)	Function (Language cultivation)
Society (Status planning)	1. Selection (Decision procedures)	3. Implementation (Educational spread)
	a. Identification of problem	a. Correction procedures
	b. Allocation of norms	b. Evaluation
Language (Corpus planning)	2. Codification (Standardization procedures)	4. Elaboration (Functional development)
	a. Graphization	a. Terminological modernization
	b. Grammatication	b. Stylistic development
	c. Lexication	

(Source: Haugen, 1983:275)

As it has been clearly indicated in the Haugen's (1983:275) revised language planning model, selection decision is all about political and social whereas codification is purely linguistics. Codification is all about preparing prescriptive orthography, grammar and dictionary. As to Hauge (1983:274), implementation is the processes of spreading the new norm, which are more consciously applied in school by teachers and textbooks, media and eventually become self-administered through reading and general social acculturation and their evaluation as an important feedback mechanism in the implementation process to assess the success. The author also states that elaboration is the continued implementation of a norm to meet the functions of a modern world.

2.1.1.1 Why language planning?

Language planning decisions, generally, may arise in response to sociopolitical needs. From Ager's (2001:44) point of view, when language planning decision refers to the standardization of a language, its purpose will be to typically attempt to reduce linguistic variation. It may include a decision where a single variety of a language is declared "standard" to promote linguistic unity in a country where divergent dialects exist. The author further states the instance of China. Although many dialects of Chinese exist, the promotion of a single variety as the national language contributes to a sense of national unity. But, here, promotion of a single variety is not the only option to standard. In some other cases, like the case of Oromo, composite of dialects approach is being followed in the process of standardizing the language.

Ager (2001:45) notes that LP and policy issues are motivated by political factors which include identity, insecurity, ideology and image; and social factors such as inequality, integration and economic instrumentality. These issues can be turned into goals that language planners may seek to achieve. For instance, the issue of identity may be a serious concern in a multilingual society like Ethiopia. Hence, there may be a need for a language policy that is all encompassing for the various speech communities in a given country. The author further claims that language may also play a significant role for social inequality. If it is planned, it is possible to reduce ethnic and political dissatisfaction on the part of the people by creating an environment based on the principle of language equality. Ideology is one such goal that motivates the planning of language. This goal may mainly be associated with the elite and power (Bamgbose 1991:178). The other goals of language planning can include conservation of the dialect forms, development of the dialects, expansion of the number of speakers of the dialects (first language or second language speakers), integration of the specific dialects (or multiple dialects) in specific societal domains, provision of services in the dialects and so forth.

Wardhaugh (2006:358) mentions four typical ideologies that may motivate actual decision-making in LP in a particular society: linguistic assimilation, linguistic pluralism, vernacularization, and internationalism. Linguistic assimilation is the belief that everyone should learn the dominant language of the state. On the other hand, the nation-state ideology with its

tendency to adopt a national and official language for a state changed Europe's geopolitical map especially since the nineteenth century, and causes the endangerment of less socially powerful languages (Wardhaugh 2006:358; Fishman 1989:97). Scholars like Fishman (1973:42) and Spolsky (2004:58) concur that the development of nationalism is the source of the need for LP.

Nationalism often favors linguistic assimilation to make sure that every member of a speech community is able to use the dominant language (Cobarrubias 1983:63). This results in transferring prestige to the dominant language to assert its superiority and to suppress the minority languages. It also triggers language shift and ultimately language death. In contrast, linguistic pluralism promotes the coexistence of different language groups and their right to maintain and cultivate their languages on an equitable basis (Cobarrubias 1983:65). It gives rise to territorial or individual-based policies or any combination of the two (Wardhaugh 1992:348). The recognition of more than one language can be complete or partial, so that all or only some aspects of life can be conducted in more than one language in that society.

Vernacularization is the elaboration of a local language and its adoption as an official language. Local languages are commonly modernized and officially recognized alongside an international language of wider communication (Cobarrubias 1983:66). On the other hand, internationalism mirrors the ideology of adopting a non-indigenous language of wider communication as an official language or an instructional medium. According to Cooper (1989:35), the objectives for LP include consumer protection, scientific exchanges, national integration, political control, economic development, the creation of new elites or the maintenance of old ones, the pacification or co-option of minority groups and mass mobilization of national or political movements. Ferguson (2006:101) also claims that LP can serve as a means for empowering groups and individuals to create and strengthen national unity, and for maximizing educational and economic development. The author further states that, contrary to the above, LP can also be used to maintain and perpetuate oppression, social-class discrimination and social and educational inequity.

Nahir (1984:295) identified goals of LP activities around the world, which are generally concerned with language choice (officialization, nationalization, standardization, vernacularization, revival, spread, maintenance, and inter-lingual communication) and language development (purification, reform, modernization, terminology, unification, stylistics, simplification and graphization). Clayton (2006:19) states that the above language-planning goals should be taken into consideration by government and non-governmental organizations as well as by individuals. As a result, LP influences a language use at a family, government, business and other larger arealevels.

2.1.1.2 Types of language planning

Under this sub topic, types of language planning and various sub activities under language planning are discussed. Accordingly, status planning, corpus planning, acquisition planning and prestige planning are discussed.

2.1.1.2.1 Status Planning

Status planning deals with language policy and its implementation as well as the selection of languages used for official purposes and education (Kaplan & Baldauf 1997:30). They argue that status planning involves the allocation of languages to different societal domains, such as the official sphere, education, business, media, etc. The explicit proclamation of a language as the official medium of communication naturally enhances its importance to a significant extent, but the introduction of a particular language in schools, for example, can have far-reaching consequences. It also deals with the social and political implications of choosing a language and with such matters as language attitudes, identity, and minority rights. In other words, status planning refers to authoritative decisions to use a language for important official function and thereby enhancing its social prestige (Webb & Kembo-Sure 2000:8). According to Kaplan & Baldauf (1997:30), status planning is concerned with attempts to modify the environment in which a language is used.

Cooper (1989:99) enumerates the following environments in which a language can acquire new functions during status planning efforts: official, provincial, wider communication, international, capital, school subject, literary and religious. Thus, status planning refers to increasing the use of

a language within a society across various domains and institutions. Status planning is concerned with changes in the functions of languages by using a sociolinguistic concepts and information to implement them (Cooper 1989:38). Therefore, issues of status planning include designation of the medium of instruction in schools and decisions regarding multilingual language use. According to Kloss (1986:140), it is not the scientist who decides on the status of a language, but its speakers by their active language use in political, educational, and cultural domains. The political domain refers to the language for administrative purposes at regional, national and international levels. The educational domain includes the choice of a language as a medium of instruction for specific regions, diverse age groups, various educational levels, and diverse school subjects.

The cultural domain refers to the language used in radio and TV, newspapers, literature, and cinema. Crystal (2003:95) is of the opinion that status planning is mainly concerned with the social and political implications of choosing a language. Crystal claims that when a government recognizes the language of a subordinate minority as a statutory or official language, it grants symbolic recognition of that group's right to maintain its distinctiveness. Therefore, the specification of legislative languages is the manipulation of political symbols for the maintenance of ruling elites in most cases. Language is essential to thinking and learning; children must be able to learn in the language. From this point of view, English and Amharic were given favor for language of learning during before the Derg regime (Fishman 1973:25). But currently, there is strong argument in favor of mother tongue as languages of learning in Ethiopia since such approaches to education have been shown to be seriously flawed, and there is comprehensive and convincing evidence which demonstrates the linguistic, academic and social advantages of mother tongue education and bilingual schools to ensure that all pupils have equal access to the school and the curriculum (Crystal 2003:95).

According to Fishman & Cobarrubias (1983:51) the status of a language is a concept that is relative to language functions. It is also relative to other languages and their suitability and eligibility to perform certain functions in a given speech community. They claim that factors which determine the status of a language include the number of its speakers, their wealth and the importance of what they produce, their social cohesiveness, and their acceptance by others.

2.1.1.2.2 Corpus planning

Corpus planning refers to such activities as the production of grammars and dictionaries, the design of orthographies, the choice of script, spelling reforms, the production of primers and readers (Bamgbose 1989: 10). To Bamgbose (1989:10), corpus planning activities includes design of orthography or reform of the existing spelling system; standardization/ harmonization of spelling/word forms; determination of word pronunciation. In addition, choice of script; terminology, vocabulary expansion; change in grammatical structure; creation of simplified registers for special purposes are also activities in corpus planning. The above author further states that dialect leveling; cultivation and counseling in respect of different styles and genres; production of primers, readers, manuals inconnection with literacy training/schooling; translations of various kinds of texts like the Bible, dictionaries, grammars; fiction and other creative arts, including language; other forms of production of written and oral texts in the broad sense; creation of institutions dealing with language questions (such as language commissions) are among corpus planning activities.

Corpus planning also deals with standardization processes that includes selection, codification, language elaboration and implementation. It determines standards and norms for a language, and develops new and technical terms for educational and public sectors (Webb & Kembo-Sure 2000:16). Corpus planning activities often arise from beliefs about the adequacy of the form of a language to serve desired functions (Wardhaugh 2006:357). When a language is identified as appropriate for use in a specific situation, efforts are made to fix or modify its structure. According to Kloss (1986:138), corpus planning aims to develop the resources of a language so that it becomes an appropriate medium of communication for modern topics and forms of discourse, and is equipped with the terminology needed for use in education and administration. It is a prescriptive intervention on the forms of a language (Coulmas 2007:38).

Magwa (2008:41) also states that corpus planning is concerned with the change of language structure that involves expanding of vocabulary, devising orthography, indicating ways of pronunciation, and style of language use. In other words, corpus planning involves the development of new forms, the reform of old ones, or the selection from alternative forms in a

spoken or written code. Corpus planning stands for deliberate changes to the internal structure of a language so that it may meet specific requirements in official domains and in domains of higher education for scientific and technical discourse (Nakin, 2009:27). According to Wolff (2000:333); Kaplan & Baldauf (1997:38), corpus planning can be described as linguistic or internal planning that is geared towards establishing spelling norms, setting norms of grammar, and expanding the lexicon. Furthermore, Kaplan & Baldauf (1997:38) mention different aspects of corpus planning like orthographic innovation (including design, harmonization, script changes, and spelling reforms), pronunciation, and changes in language structure, vocabulary expansion, and simplification of registers, style, and the preparation of language material. For Hornberger (2006:30), corpus planning comprises lexical modernization, purification, reform, stylistic simplification, and terminology unification.

Webb and Kembo-Sure (2000:8) call the preparation of grammars and dictionaries *codification*. Codification deals with the way language norms are chosen and codified and involves the selection of a national language, reformation of the spelling system, launching of campaigns for plain and non-sexist language and introduction of literacy programs (Crystal 2003:358). Reagan (2002:390) notes that corpus planning focuses primarily on lexical development of specific languages by creating new terminology and dictionaries, changing parts of the grammatical system and develops textbooks (Bamgbose 1989:24 & Mutasa 2004:20). This means, corpus planning involves the development of a language in its totality. The three traditionally recognized types of corpus planning are graphization, standardization, and modernization.

Graphization refers to the development, selection and modification of scripts and orthographic conventions for a language. The use of writing in a speech community can have several socio-cultural effect; such as an easier transmission of material across generations, communication with larger numbers of people, and a standard against which varieties of spoken language are often compared (Ferguson 2006:8). Ferguson (2006:8) made two observations about the results of adopting a writing system. First, the use of writing adds another variety of the language to the community.

Vocabulary, grammar and phonology of a language commonly adopt characteristics in the written form those are distinct from the spoken variety. Second, it is often believed that written language is 'real' language, which ignores the possibility that the written language may have been based on a variety of the spoken language (Trudgill & Chambers 1980:63). Writing systems provide the basis on which literacy materials are established. They have the potential to reduce linguistic variation in a language community and are often associated with the transformation of an oral language with a written language and modernization (Coulmas 1998:64; Wurm 1994:211). Mabule (2011:37) states that there are several aspects to consider in graphization such as orthographic conventions, representing allophones with separate symbols, alphabets versus syllabaries, political and/or social acceptance of the alphabets/ syllabaries, learnability of the new alphabets/syllabaries, and their transfer between languages.

Modernization, which is usually a concern of majority and national languages, refers to the creation of terms for new concepts (Ferguson 2006:102). According to Webb & Kembo-Sure (2000:81), modernization consists of borrowing and adapting new technical terms from the language in which they were invented. Modernization is a form of LP that occurs when a language expands its resources to meet new functions like a shift in status when there is a change in language policy in education (Gottlieb 2005:25). The most significant force in it is the expansion of the lexicon, which allows the language to discuss modern topics. As to the above author, language planners generally focus on creating new lists and glossaries to describe new technical terms, but it is also necessary to ensure that the new terms are consistently used by the appropriate sectors within a society. Rapid lexical expansion is aided by the use of new terms in textbooks and professional publications. Issues of linguistic purism often play a significant role in lexical expansion, but technical vocabulary can be effective within a language, regardless of whether it comes from the language's own process of word formation or from borrowing from another language (Edwards 2004:181).

Cooper (1989:125) stated graphization, standardization, and modernization as elements of corpus planning. The concepts of codification and elaboration are usually discussed under corpus planning and these concepts can be subsumed under the sections of standardization and modernization, respectively. In addition, Cooper (1989:154) suggested *renovation* as the fourth

major category, while modernization focuses on the ways of established language forms to serve a new function; whereas renovation aims at finding new forms to serve the existing functions in innovative ways.

2.1.1.2.3 Acquisition planning

Acquisition planning is the type of LP which deals with challenges in language revitalization, maintenance, shift, and the teaching of second and foreign languages (Ferguson 2006:95). Many scholars restrict acquisition planning to a teaching of major second as well as foreign languages (Lambert 2000:45; Cooper 1989:157). However, acquisition planning should not be limited to these languages but also take into considerations the teaching of local and minority languages (Adrey 2009:4). To Kaplan and Baldauf (1997:125), it is very pivotal to recognize that status, corpus and acquisition planning do not take place in isolation but, are integrated into LP processes in which the statuses of languages are evaluated, corpuses are revised and the changes are introduced to society at various levels through education systems. The authors list five steps which form the major considerations for a language-in-education policy and implementation: curriculum, personnel, materials, community, and evaluation. Changes in the language-in-education policy are achieved through the development of new curricula, and the use of new teaching methods and materials. The curriculum should take into consideration the demographic composition of a society and reflect its socio-cultural diversity by integrating some local and other cultures. Acquisition planning reflects the demands of an increasingly multilingual society or in second and foreign language learning (Antia 2000:4).

Language policy involves decisions concerning the teaching and use of language, and their careful formulation by those empowered to do so (Adrey 2009:106). As to the author, language spread is thought of as promoting the acquisition of a new language or as promoting a variety of a particular language as the standard. In multilingual situations, this could involve raising the status of minority languages. Education is arguably one of the most important aspect of language planning. Adrey (2009:106) further states that through education policies, corpus planning can be reinforced as children are taught the prescribed form of a language. Status planning can be enhanced when the importance of a particular language is emphasized by its role in the curriculum; and acquisition planning is centrally realized through education programs. Since LP

involves language teaching, Cooper (1989:33) argues that those aspects of planning that are directed towards increasing the number of users deserve a separate analytic category.

Thus, educational function of language is one of the most critical aspects in LP and policy. Cooper (1989:108) describes it as “the function of a language as a medium of primary or secondary education, either regionally or nationally”. This function is not limited to primary and secondary education, but also applicable to tertiary levels of education as well. Fasold (1984:76) argues that a language requires three attributes in order to fulfill its educational function, namely (i) it must be understood by learners, (ii) there must be teaching resources, i.e. textbooks as well as teachers able to teach in the language, and (iii) it must be standardized. Fasold (1984:76) observes that in a situation where a community has a language that fulfills these three attributes is standardized and has teaching resources is very often selected by the students come to understand. In fact, efforts to upgrade the status of the mother tongues of such minority students should take priority over efforts to reduce the educational disadvantages associated with the use of foreign languages for teaching (Fishman 1989:327). Acquisition planning should include activities such as teachers’ selection and training, besides activities like content, time and resource allocation, and selection of teaching materials. Evaluation at each stage of acquisition planning is very significant to get feedback (Baldauf & Ingram 2001:412).

2.1.1.3. Language standardization

Standardization is one of the elements of language planning. According to Webb & Kembo-Sure (2000:18), language standardization is the process by which an authoritative language body prescribes how a language should be written, how its sounds should be pronounced, how its words should be spelt, which words are acceptable in formal situations, and what grammatical constructions are appropriate. The issue of standardization is crucial for language use as a set of practical guidelines for appropriate public verbal behavior. Language standardization involves status planning, when it refers to the selection of a single variety of a language as the standard language, and corpus planning, when it refers to the codification of a language in a unified variety. It is a process that any language needs to undergo in order to develop its vocabulary and writing system so that it becomes capable of expressing whatever functions the speakers need to express. Any language variety or dialect is capable of developing attributes essential for

standardization. Since the issue of language standardization is the main focus of the present study, detailed discussion on the matter is presented in the next sections.

Deumert (2004:45) defines standardization as a natural development of a standard language in a speech community or an attempt by a community to impose one dialect as standard. But, Hudson (1980:32) argues that it is a direct and deliberate intervention by society to create a standard language where before there were only dialects as nonstandard. The process of standardization contains four basic steps: (i) selection of the norm, (ii) codification of the form, (iii) elaboration of the functions, and acceptance of the variety by the community (Deumert 2004:45). Joseph (1984:9) & Coulmas (1997:82) are of the opinion that standardization is a process of converting varieties into a standard by fixing and regulating its spelling, grammar and vocabulary. Webb & Kembo-Sure (2000: 18) define standardization as the process by which an authoritative language body prescribes how a language should be written, how its sounds should be pronounced, how its words should be spelt, which words should be accepted in a formal situation and what the appropriate grammatical constructions of the language are. The authors also state that standardization is not an inherent, but rather an acquired or deliberately and artificially imposed characteristic. Standard languages do not arise via a natural course of linguistic evolution or suddenly spring into existence; they are created by conscious and deliberate planning (Webb & Kembo-Sure (2000:18).

Standardization is, therefore, a process by which a dialect(s) changes in status, function and form to become a standard. The process deals with linguistic aspects such as morphology (form and structure of words), phonology (sounds of a language), syntax (sentence structure) and spelling. Standard languages are the products of the process of standardizations. They are regarded as a planned creation because their development process can be traced. To Romaine (1994:84) standardization is not an inherent, but rather an acquired or deliberately and artificially imposed characteristic. From the author's perspective, standard languages do not arise via a natural course of linguistic evolution or suddenly spring into existence. They are created by conscious and deliberate planning.

Standardization is needed for the development of materials such as standardized and harmonized grammars, dictionaries and other educational materials for the use in mother tongue education in schools (Hudson 1980:34). In order to standardize a language, first its various dialects need to be identified and analyzed, then the core dialect is chosen from which the standard form is constructed through the process of harmonization of the dialects, usually in the form of a standard grammar book, dictionary and literary publications (Ager 1996:35). In practice, standardization generally entails increasing the uniformity of the norm, as well as its codification (Ferguson 2006:65). It is the process by which one variety of a language takes priority over other social and regional dialects. This variety becomes a supra-dialectal form of the language (Wardhaugh 2006:27). Gottlieb (2005:19) states that the standard is commonly spoken by the most powerful social group within the society which is imposed upon the less powerful groups. This often reinforces the dominance of the powerful social group and makes the standard norm necessary for socio-economic mobility (Wardhaugh 2006:29).

Language standardization is necessary for many reasons. It allows a language to serve as the common means of communication within a community (Wardhaugh 2002:34). In addition, a mutually comprehensible variety unifies the members of a community. Wardhaugh (2002:78) mentions more positive aspects of a standard language, like its function as a symbol for a common identity, or the prestige which it can give to its speakers. Standardization also makes it easier to teach a language in schools.

Standard language services as an instrumental, communicative and symbolic function. Standard language is instrumental insofar as it is learned and used by the member of a nation or region (Wardhaugh 2002:78). A standard becomes tool or instrument by which information spreads through educational, governmental, religious, cultural and media institutions and by which these different mechanisms of standardization uphold common national idiom. Joseph (1987:19) also states that standard languages are acquired largely, even primarily, through instruction, correction, imitation, assimilation, and acculturation.

Regarding the instrumental function of standard language Mirloy & Mirloy (1999:19) assert that the whole notion of standardization is bound up with the aim of functional efficiency of the

language and what is essential is that everyone should use and understand the language in the same way with the minimum of misunderstanding and maximum of efficiency. The other function of a standard language is communication function which is a vehicle through which normal social interaction are conducted. To Cameron (1995:23-27), communication is a maximum of language prescriptivism and needs a standard language. As a standard language has to be learned, it has to be accessed through education and other media via which the harmonized language is diffused.

A standard language functions a symbolic role. Nations express their belongingness and loyalty to a community, not so much by what they said; but their use of specific language variety - standard language (Paffey 2008:45). The author goes on explaining that the symbolic function of a standard language has its root in linguistic nationalism that considers language, nation and citizenship to be historically and spiritually congruent. Hence, from this point of view, a standard language does not merely serve the function of communicative efficiency, but is a commodity which symbolizes one's belongingness to the overarching nation or state or region. Language is employed to unite and to distinguish nations. It also shows ethnic, regional, gender, class, cultural features with the aim of projecting self and determining the other person.

Unfortunately, the benefits of standardizing a language are accompanied by several negative effects. Wardhaugh (2006:34) wrote, "... choosing one vernacular as a norm means favoring those who speak that variety. It also diminishes other varieties and possible competing norms, and those who use those varieties." This creates a sense of elitism among the speakers of the standard and a feeling of inferiority in those who continue to use the non-standard form. In situations where two stable varieties co-exist, Fishman (1970:30) observed that one speech variety (the standard) is usually associated with status, high culture, and aspiration towards social mobility, while the second variety (non-standard) is typically connected with solidarity, comradeship, and intimacy with a low status group. This tends to create an inferiority complex among the low status speakers. It also forces parents who wish for children to be accepted by the dominant group to conform to the standard, possibly neglecting the language and culture they were born into.

Lippi-Green (1997:59) believes that the concept of a standard language is a myth, created and perpetuated by social groups with the most power in a given society at the expense of all other language groups. She notes that individuals act for a larger social group takes it upon themselves to control and limit spoken language variation. Kamwangamalu (1997:237) says that a standard form is usually associated with prestige as it provides a unified means of communication and an institutionalized norm which can be used in mass media, education, teaching the language to foreigners etc. Trudgill & Chambers (1980:161) argue that standardization is necessary to assist communication, to arrive at an agreed orthography and to provide a unified form for school textbooks. Schiffman (1996:166) explains that the standard language in a community is related to the prestige of the languages spoken in the community.

In general, greater prestige tends to be attached to the notion of the standard since it can function in higher domains, and has a written form. The acquisition and adoption of the standard must be distinguished from the equally important acquisition of the technical ability to write. As noted by Milroy (2001:539), standardization is one of the more difficult requirements of a reference norm so that it can accommodate and adjust to new needs and changes. In order to fulfill the functions of modern scientific, technological, administrative and educational communication, a standard language has to be intellectualized through the modernization and elaboration of its vocabulary.

2.1.1.3.1 Development of standard language

The principal characteristics of a standard are that it begins as one of many, linguistically, equal spoken dialects, it has an alphabet and consequently a written form. This codified and written form of a speech is then promoted from a dialect to a language (Haugen 1972:97). Then, it is superimposed and taught by national institutions. In doing so, it is viewed not only as a function tool; but also as an icon of national identity (Paffey 2008:44). How can a language be standardized is a very important question to be addressed.

The objective of selecting the norm is to take a language variety that will be developed as the standard language. Selection can be done in various ways. For example, one dialect may be selected from a number of related dialects of the same language, and developed into a standard, as in the case with Sepitshu in South Africa (Malimabe 1990). Another way is to construct a

variety by "combining" features of a number of related dialects of the same language to form one variety. Problems may encounter with both the composite way and the single dialect way. Wardhaugh (2002:31) observes that selection may prove difficult, because choosing one vernacular as a norm means favoring those who speak that variety. A composite of dialects may also be regarded as favoring some dialect features and marginalizing others, and those speakers who feel disadvantaged may complain. Members of the strongest dialect community may also object to learning and using features from other dialect communities. The issue of language standardization can be controversial and sensitive. An authoritative language body must try to be fair, but everyone may not be happy with the decision made by the authority. Choosing the lesser evil is better.

Haugen (1972) suggests taxonomy of four stages through which languages pass in order to be standardized. These are selection, codification, elaboration and acceptance. The stages of standardization necessarily begin with the choice of which variety will become the common language. It is extremely rare to find a nation or state in which every citizen speaks the same vernacular. Selection involves choosing a variety(s) for specific purposes frequently associated with official status or normal roles. In other words, norm selection is the making of the official policy regarding the language. Standardization involves at least two distinct language strategies. One requires the elaboration and adoption of one variety among others; while the other consist the creation of a new variety which is composed of some main dialects. Therefore, decisions to base a national linguistic unity around any one language variety will actually lead to a significant degree of exclusion of other varieties. Haugen (1972:109) states that choosing any one vernacular language as a norm means favoring the group of people speaking that variety. This vernacular language gives its speakers prestige as a norm bearers and a head start in the race for power and position. If recognized elite already exists with a characteristic vernacular, its norms will almost an inevitable will prevails (Haugen 1972:109). To the contrary, where there are socially coordinate groups of people with the community, usually distributed regionally or tribally, the choice of any one will meet with resistance from the rest.

Hence, while a selection of vernacular to be a standard may at first appear to be a matter of endorsing one variety and promoting its use, it also means, by inference, all other varieties are

deselected and implicitly or explicitly excluded from the status accorded to the standard. To Lippi-Green (1997:173) the concept of a standard language is concerned with not so much with the choice of one possible variant, but with the elimination of socially unacceptable differences.

The second stage of standardizing a language, to Haugen (1972:107), is a codification which encompasses efforts to fix the selected language variety, firstly through alphabetization, if this does not already exist; through securing conventions on orthography, lexicon and grammar, all with the aim of achieving minimal variation in form. Codification refers to an explicit statement of the standard language form via dictionaries, grammars, and punctuation and pronunciation guides (Appel & Muysken 1987:51). Codification typically develops and formalizes a set of linguistic norms for a written language (Kaplan & Baldauf 1997:39). This stage is related to the stabilization of the norms and refers to a description of features in normative grammar. Codification presupposes norm selection and is related to a standardization process. Crystal (1980:243) describes the concept norm as "a standard practice in speech or writing". On the other hand, Bartsch (1985:62) defines a norm as: 'a prescription that has been accepted and internalized by the society as a standard by itself, guiding one's own behavior as 'wrong' or 'right''. This shows that norm is a standard or a prescription used as "right or 'wrong' way of using a language. A norm is also practical guidelines for appropriate linguistic behavior. They are context-sensitive rules for language use and provide a cultural definition of desired behavior.

From the above author's perception about a norm, it can be deduced that norms guide the society in performing certain functions in an appropriate and desired manner; in this instance, they are guidelines on the spelling of words and their pronunciation. There are two types of norms; namely formal and informal. The informal norms generally arise spontaneously. They exist naturally within a speech community and constituent monitored social practice in the society (Bartsch 1985:62). The informal norms apply to a spoken language and are passed on from the older generation to the younger generation. The informal type of norm contains some expectation about social behavior, thus, what the society regards as being good behavior.

The formal norms, on the other hand, can be said to be externally determined. Language authorities develop these norms for the standard variety of a language. Norms of this nature are

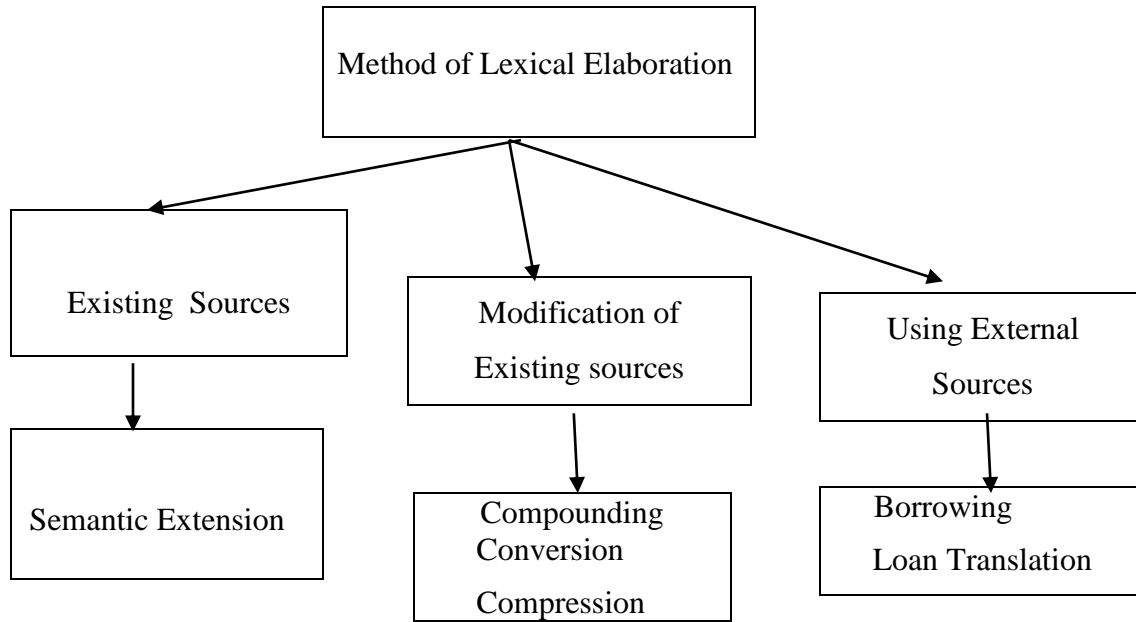
created for a purpose. Bartsch (1985:62) points out that these norms are consciously developed and are obligatory to a certain degree. This type of norm is not static; what is regarded as a norm today may not be a norm tomorrow.

The third stage is elaboration which is a way of making a language capable of expressing modern new concepts. In this stage, the variety is developed and prepared for use in various formal domains and all functions associated with the government. These functions include usage in mass media, religious organizations, parliament, courts of law, and education and in literature (Hudson 1980:32; Wardhaugh 1992:30). In the elaboration stage, the acceptance of the standard variety is also promoted. Romaine (1994:13) pointed out that elaboration stage leads to the maximal variation in function. This implies that the standard variety should have a wider range of uses in comparison to other varieties of the same language. The functions must meet the needs of the community at large. This stage involves the expansion of the language functions and assignment of new codes, such as scientific and technological. Language modernization is one of the most common activities requiring elaboration. Elaboration involves the terminological and stylistic development of a codified language to meet the ongoing communicative demands of modern life and technology (Deumert 2009:379). Cooper (1989:125-126), states that elaboration is the development of inter-translatability of a language with other languages in a range of topics and forms of discourse characteristics of industrialized, secularized, structurally differentiated, 'modern' societies. On the other hand, Wolff (2000:337) stresses the need for elaboration of vocabulary for commercial, professional, and scientific domains as knowledge and technology progress. Similarly, Ferguson (1996:46) highlights that elaboration enriches languages with the specialized vocabularies, and forms of discourse in relation to various functions the language must accomplish in a modern society.

Although they are not the only concepts to be taken, Haugen's (1983) language planning and standardization, Sager's (1990) mechanisms of lexical elaboration, word or term formation strategies, Cabré's (1999) neologisms (techniques of creating new words), ISO 704 (1999) and Felber's (1984) standardization principles are mainly considered in this dissertation.

Sager (1990:71) states that methods of lexical elaboration are conceptualized into three basic types, as shown in Figure 1: (i) use of existing sources, (ii) modification of existing sources, and (iii) use of external sources. Those stated under (i) and (ii) comprise language internal resources, whereas (iii) represents language external resources of lexical elaboration technique.

Figure 1: Methods of lexical elaboration



Source: Modified based on Sager (1990:71)

As it is stated above, Sager (1990:71) states that there are three main approaches to the development of new designation so as to represent concepts: These are (1) using existing resources which are extensions of existing terms to express new concept; for instance, the use of simile or naming of concept in analogy to another, is familiar one. (2) modifying existing resources as terms can be formed by derivation (or affixation), compounding, conversion and compression are the most common method of designation of new concepts by modifying existing resources and (3) creating new linguistic elements or neologisms. The re-use of a special designation in one field to another can also be used for the naming of different concepts. On the other hand, neologism is a method of term formation in which new concepts, particularly in

science and technology are formed that is totally new creation or a borrowing from other languages (Sager 1990:79).

Borrowing and loan-translation are among productive ways of creating new terms both of which are external source for the expansion of vocabulary in a language (Cabr  1999:94). As to Myers-Scotton (2006:211), borrowing is defined as incorporating an item or words from one language and it is of two types; cultural and core borrowing. The former, is a filling of gaps in the recipient language's store of words as they stand for objects or concepts new to the language's culture. The latter is borrowing words which duplicate elements that the recipient language already has in its word store. Borrowing can be modified or unmodified. Modification is adaptation of a word into the phonetic and a grammatical system of the recipient language and unmodified borrowing is borrowing a word without any modification. A transdisciplinary and translingual are internal and the adoption of terms or concepts from one language to another respectively (ISO 704 1999).

The fourth stage of language standardization, according to Haugen (1983), is acceptance (implementation). The selected variety needs to be accepted by the community at large, so that it can serve as a communicative tool across different groups (Hudson 1980:32). The acceptance entails an agreement that has been achieved with regards to words that can be used or left out. The cooperation and involvement of the society is essential. A problem that might be encountered at this stage is that people may not agree with the selected form for some reason. For instance, where a single variety has been selected, people may think that their vernacular is being "despised" and not seen as important.

Lane (2018:111) states that developing a written standard always entails making choices of what to include and what to leave out, which ultimately translates to choices about who to include and who to leave out. In this case, one way in which the inclusion of standard and exclusion of nonstandard languages is achieved is through efforts to enlarge the body of standard language users. In order to strengthen the position of one of particular standardized dialect, members of a community must accept its usage in common life.

Standardization is concerned with linguistic forms (corpus planning, i.e. selection and codification) as well as the social and communicative functions of language (status planning, i.e. implementation and elaboration) (Haugen 1966:5). Language standardization emphasizes the desirability of uniformity and correctness in language use, the primacy of writing and the very idea of a national language as the only legitimate language of the speech community. Linguistically-oriented approaches to language standardization have often concentrated on the identification of the regional and/or social dialects which form the phonological, morphological and syntactic basis of a standard language (Deumert 2003b:6). As to (Deumert 2003b:6), however, standard languages which are based on a single dialectal source (mono centric selection) are rare, and most standard language histories have been shaped by dialect leveling. The majority of standard languages are, thus, composite varieties characterized by multiple selections; that is, the complex recombination of features from various dialects and varieties (polycentric selection). In other words, standard languages have ‘multiple ancestors’ and their history is shaped by various types of language contact which are due to dialect convergence as well as spoken/written language contact; Even by the time a norm has been codified and elaborated by its users, it has become virtually impossible to identify its base.

From a variationist perspective, standardization can be conceptualized as a movement towards linguistic uniformity through a competition-selection process: certain variants or linguistic habits are selected as part of the standard norm and are generalized to new linguistic and communicative contexts (Lodge 1993: 95). Linguistic variability and heterogeneity, on the other hand, become indexical of non-standard varieties. As to Lodge (1993: 95), variability, however, does not imply the absence of norms. The author also claims non-standard varieties are characterized by a multiplicity of highly context-specific, particularistic norms which have emerged in response to the local needs of the loosely networked social groups which make up the speech community. The norms of standard languages, on the other hand, are universal and show little contextual, geographical or social variation. Both variation and standardization contribute to the formation of sociolinguistic groups: while the maintenance of variation marks social, ethnic and regional differences within the larger speech community. To the contrary, standardization promotes social and political unification and a common identity.

2.1.1.3.2. Characteristics of standardized language

A standardized language is considered necessary to reinforce the potential for communicating and spreading knowledge and information, as well as for education. However, the main reason is that a standardized language could be a symbolic resource for getting speakers to think of themselves as members of a shared ethnic group (Geeraerts 2003:5). Standardized language seeks a collective symbolic dimension and implies a political use of the language. A standard language is linked to the notion of language as an institution, inflexible and rigid.

Geeraerts (2003:5) states that the most conspicuous feature is the generality of standard languages. Standard languages, in contrast with dialects and other restricted languages are general in three different ways. They are geographically general, in the sense that they overarch the more restricted areas of application of dialects. Further, they are socially general because they constitute a common language that is not the property of a single social group but that is available to all. Finally, they are thematically universal in the sense that they are equipped to deal with any semantic domain or any linguistic function. More advanced domains of experience in particular (like science or high culture) fall outside the range of local dialects. Because of their generality, standard languages have additional features (Geeraerts 2003:5).

Standard languages are supposed to be a neutral medium, with a mediating function, in an almost philosophical sense of 'mediation'. The author further states that standard languages transcend social differences. They ensure that men and women from all walks of life and from all corners of the nation can communicate freely. In that sense, they are a media of participation and emancipation. Because of their neutrality and because of their functional generality, standard languages are keys to the world of learning and higher culture: functional domain par excellence for standard language use (or, reversing the perspective, functional domains that cannot be accessed on the basis of dialect knowledge alone). Perhaps even more importantly, standard languages are supposed to contribute to political participation.

The possibility of free communication is a feature of a democratic political organization. If then, linguistic standardization contributes to mutual understanding and free communication. It is a

factor of political emancipation just as it is a factor of social emancipation when it contributes to the spreading of culture and education.

2.1.1.3.3 Principles and methods of standardization

Since standardization is a deliberate and an intentional intervention by concerned body of the government or by society to create a standard language from non-standard varieties, it is obvious that there are principles and methods to do it. The principles and methods of standardization are discussed in the following subsections.

2.1.1.3.3.1 Principles of language standardization

The main focus of standardization is to minimize or to suppress variation of written language so as to avoid miscommunications. To standardize a language, the concerned body has to produce dictionaries, grammar books, spellers, and style manuals due to the fact that the main target of standardization is on the written rules of a language (Cooper 1989:144-145). Creating new terms or standardizing existing terminologies needs principles to guide. From this point of view, Cabré (1999:200) states that in the standardization of a language, sociolinguistic factors (usage, language policy, user needs, etc.), psycholinguistic factors (customs, morals, aesthetics, etc.) and formal linguistic criteria (well-formedness, possibilities for derivation, morphological motivation, etc.) should be considered.

Transparency, consistency, appropriateness, linguistic economy, derivability, linguistic correctness and preference for native language speakers are standardization principles stated by ISO 704 (1999:25) and Felber (1984:179-186). In order to avoid miscommunication and to maintain the appropriateness, terms formed should follow established patterns of meaning. From ISO 704 (1999:23) point of view, “A *term* is considered transparent when a concept it designate can be inferred, at least partially, without a *definition*”. It can be inferred that transparency is the visibility of meaning in the morphology of a language. The term transparency implies an explicit relation to a certain concept. On the other hand, consistency is all about avoidance of arbitrariness during the development of terms. To attain consistency, a terminological system should be structuralized. Appropriateness is one of the principles that is important during term formation. It is a term closely connected to its connotation and its use within a particular

language. Appropriateness requires a term to be neutral; avoiding connotations. The structure of a term should not lead to ambiguity and must exclude confusion.

Another important principle to be considered in standardization is linguistic economy. Conciseness implies that a term has to be as short as possible to be easily retrieved and reproduced. As it is stated in ISO 704 (1999:26), the requirement for conciseness often conflicts with that for accuracy. The greater the numbers of characters included in a term, the greater the precision and transparency of the term. However, increasing the number of characters often makes a term too long and inconvenient to use. Terms shall be as shorter as possible to be easily understood and be economical. Another principle of term formation is derivability. It refers to the productiveness of term during the process of term formation. With linguistic correctness, it means a term shall conform to the morphological, morphosyntactic and phonological norms of the language in question (ISO 704 1999:27). There are still other principles of standardization like preference and acceptance by native language and the existence of a concerned body to foster such issue. Standardized language or language variety will assume the status of a standard language only when people come to an agreement to accept it as a standard language by developing favorable attitude towards its use (ISO 704 1999:27).

2.1.1.3.3.2 Methods of standardization

In some countries several agencies could be assigned with the power of standardizing a language. This is due to the fact that language planners and a society itself feel that language is deficient because it could not be used as such to carry out all the conceivable functions which arise due to the heterogeneous growth of a society (Cabré 1999:205). In certain countries and societies, some agencies are voluntarily established to undertake the work of the planning of a language. In some cases, elite or institutions belonging to a society come forward on their own to take after the activity of language planning.

From Haugen's (1966:178) point of view, there are a number of agencies are involved in the process of standardizing any language. By individuals it means people who are concerned with language, in one way or the other, may be grammarians, lexicographers, literary personalities, etc. These people normally initiate the process of standardization in some form or the other.

Haugen states that it is after the work of the initiation of the language standardization stakeholders such as organization of teachers, literary bodies, academicians and mass media that people then come forward to develop the processes involved in standardization. Textbook developer body, translation body, media agencies and related bodies are constantly involved in the process of language standardization.

The major language standardizing agent is the government. Government, in accordance with the needs of the people from different bodies for the preparation of a standard textbook for the coinage of technical terms, making of dictionaries, glossaries, etc. (Desalegn 2012:10). Haugen (1966:178) argues that governments have the advantages over private individuals in having control over the school system. Supplementary to implanting the standardization role in the schools, the government has the power to organize, guide, control, etc. institutions like language academy/language committee, the mass media, administration and the judiciary. A language academy or in some country language standardization committee refers to a responsible body for the harmonization of languages (Sadembouo 1988:5).

It is a language and cultural organization whose objective is to promote and orient the development of language. This body is responsible for proposing a standard form and its implementation. In most of the cases it is established and sponsored by government constitutionally to serve as the highest authority. It checks, controls, guides, etc., the process of standardization. Such an authoritative body is usually known as an academy in many European countries, but language committees in many African countries (Sadembouo 1988:10). In such countries, the authoritative bodies will facilitate the creation of terms through guidelines devised for the standardization policy of languages and coordinate their work by following guidelines or written criteria which are usually suggested by a central standardization body (Sadembouo 1988:10). This central body is responsible for maintaining consistency of the existing designations and creating new words. The central body usually follows written criteria for the formation of terms and the treatment of borrowings and loan translations from other languages (Cabr  1999:201). In addition, the central standardization body should also comprise different subject specialists and linguists or language experts so as to facilitate the standardization process (Samuel 2017:32).

There are various methods of elaboration such as derivation, composition, redefinition of existing terms, borrowing and compression. Redefinition of existing terms is often used in a relatively new area of study where the conceptual system itself may be unstable. Standardization is also possible by exploiting the derivational nature of the language, i.e. general or special language (Sager 1990:120). Terms can also be developed using a composition of the existing words and through borrowing from other languages. Speakers of the language can create various terms through guiding themselves using linguistic rules of the language. In this regard, Cabré (1999:94) states various types of rules for term formation:

a. derivational rules using a prefix

[prefix + [base]x] (x)(y)

b. derivational rules using a suffix

[[base]x + suffix] (x)(y)

c. compounding rules

[[base]x # [base]y](x)(y)

d. conversion rules

[base]x → [base]y

Thus, coining new words or terms are guided by a certain linguistic rule, as it can be seen above. Terms can also be created using neology which is a technique of creating new designations that is essential in special subject fields where a language has a lexical gap to express concepts (Cabré 1999:204). Cabré (1999:207) further states that neologisms can be formed through functional neologisms, semantic neologisms, formal neologisms, and borrowed neologisms. As to the author, functional neologisms refer to lexicalization of form and syntactic conversion. Neologisms in a form include compounding, phrases and shortenings (i.e. initialisms, acronyms, clippings). To the contrary, semantic neologism is the processes of broadening or narrowing or change of the meaning of the base form. Borrowed neologisms show true borrowing and loan translations.

Generally, in order to minimize variations of a written language through various methods of standardization, a language should be standardized via central national (state) body that is given a responsibility by law of the state or nation. This central national (state) body has to be guided by the guidelines outlined to standardize that language. The standardizing body shall clearly state the methods, principles and approaches of standardization of the language and there shall be guidelines printed and given to the language users.

2.1.1.3.4 Standardization and education

Education sector is one of the main sites where a normalized (uniform) codes are strictly applied. Textbooks are normally expected to be published in a standard language. The site is again very important to directly implement the codified forms of a language. In this regard, the recently ratified language policy of Ethiopia states that government institutions, educational institutions and the mass media shall have special responsibility to maintain the standards of Ethiopia's writing systems that are currently in use (MoCT 2020:23). The responsibility is not only maintaining the standard writing systems; but also the reduction of any kind of language variations in any type of texts. Therefore, the education sector experts have to compile textbooks and other learning materials using a standard Oromo via working in collaboration with the standardization committee of Oromo. All the textbooks and other learning materials have to be distributed from a single center. These authorities also have to work hand in hand to uniformly apply words of different dialects so as to maintain the stable use of the standard forms. Hence, the current working primary and secondary textbooks written in Oromo are used as a text corpus to evaluate the orthographic and lexical (de)standardization of Oromo in the textbooks (see chapter 5).

2.2 Review of studies related to standardization

The reviews of selected related studies included under this part are subdivided into three; review of selected related studies on standardization of languages of other countries, Ethiopian languages and Oromo. Various studies have been accessed; but only some studies that are directly related to standardization have been reviewed.

2.2.1 Related studies on standardization of languages of other countries

Van Huyssteen (2003) describes a practical approach to the standardization and elaboration of Zulu as a technical language in his PhD research. The study identified the poor implementation of language policy and inefficient coordination of corpus planning initiatives as the main cause for the lack of terminology in African languages in general and Zulu in particular. The finding of the study further depicts that even though African languages in general, including Zulu possess the basic tools that are vital for their development, for instance, orthographic standards,

terminological lists, dictionaries and grammar, there are deficiencies in achieving effective technical elaboration and standardization. A serious inconsistency in the formation and interpretation of orthographic rules in Zulu are also identified. Van Huyssteen (2003) claims that there is lack of documentation of word formation methods for the terminologists that could be used as a model.

Regarding the standardization of the Basque language, Hualde & Zuazo (2007:6) state that the development of standard Basque is a recent phenomenon and is undoubtedly because Basque is spoken in societies where other languages have traditionally occupied all functions beyond informal communication at the local level. Only some of the few writers who choose to write in Basque generally for religious propaganda, less commonly for literary purposes felt the need for a standard variety. Though there are some authors that published in Basque in the sixteenth and seventeenth centuries, the task of standardizing the language became explicitly difficult due to the existence of dialectal diversity (Hualde & Zuazo 2007:16). Accordingly, little progress in the standardization of the language was made until the second half of the twentieth century. The creation of standard Basque was connected to the strengthening of Basque identity sentiment in the 1950s and 1960s.

Hence, the Basque Academy was responding to this strong societal demand for education in Basque and for an expansion of the roles assigned to the Basque language, including the media and public administration. The new standard became an immediate success, in spite of an initial negative reaction in some quarters. The authors state that less than forty years after the introduction of *euskarabatua*, the domains of the Basque language have been expanded considerably, even in geographical areas where the language had been lost centuries ago (Hualde & Zuazo 2007:7). The position of standard Basque in Basque society appears secure to the extent that the Basque language itself has a secure position. In the construction of standard Basque, both the archaizing and the statistical procedures have been employed, whereas the comparative procedure, the use of reconstructed hypothetical forms, has been rejected (Hualde & Zuazo 2007:7). The authors further stated that at the present, codification in the case of Basque involves orthographic matters, establishing rules of “good usage” and deciding which dialectal lexical, morphological and syntactic variants are acceptable in the standard. From this long journey to a

standard Basque, we can note that a language can be revitalized and become a standard language if a language academy and various and individual writers as well as speakers of a language in general play their own role. In this regard, the standardization committee of Oromo, hence, can learn a lot from the standardization process of Basque in general and language academy of Basque in particular to standardize Oromo which is now in its infant stage.

Magwa (2010) in his article on language harmonization in South Africa has discussed about the orthographical standardization of the Shona language. The article seeks to discuss the harmonization of Shona language varieties; namely ChiKaranga, ChiKalanga, ChiKorekore, ChiNdau, ChiManyika, ChiBarrwe, ChiHwesa, ChiTeve and ChiZezuru. As to the author the author Current Shona orthography uses a conjunctive system of word division that was recommended by Doke in 1931. Detailed word divisions and spelling rules are found in Fortunes 1972 'A Guide to Shona Spelling'. The writing system currently being used by various Shona speakers and writers is also stated to be deficient since it does not cater for the boarder issues of Shona dialectical variations. The spelling systems being used by the Shona is purely conventional and bears very little relation to what they speak. The Article argues that a unified standard Shona writing system is achievable by way of prescribing common alphabet, common spelling, common punctuation and common word division system. It would be naïve to expect a uniform practice at the spoken level. The paper concludes by advising the speakersof the different Shona dialects that the spirit of language unification and harmonization always should be that of cooperation and not that of hegemony or conquest.

Banti & Abdirachid (2015) discusse some issues in Somali orthography.The issu es addressed by the 1960 Somali language Committee (SLC), Standardization and linguistic variation of Somali language, harmonisation of writing, the spelling of borrowed words and reference works for orthography and model of writing of Somali language are described by the authors. Banti & Abdirachid (2015) state that the main problem of the Somali language is linked to the fact that the central institution which had been in charge of its standardization collapsed with the central government of Somalia in1992. From then, and even some years before, the Academy of Sciences, Arts and Literature stopped working on the improvement of Somali language and its writing. To the authors, the Somali language Committees, especially those of 1960 and 1972,

achieved remarkable results in terms of the standardization of the language. Banti & Abdirachid (2015) believe that with the dispersion of Somalis all around the world and the emergence of a more educated generation, the development of the new technology of communication increased the usage of Somali writing. Banti & Abdirachid (2015), finally, claims that there is need of an institution, which continues the work begun by the different SLC and the academy of sciences.

2.2.2 Related studies on standardization of Ethiopian languages

Zelealem (2008) examines lexical development in Sidaama. He identified that borrowing is the most frequently occurring mechanism of lexical development followed by compounding, loan translation, semantic extension and paraphrasing in the language. Zelealem (2008) also suggests that in order to exploit internal resources of the language, searching and identifying lexicons from various varieties of the language is vital.

Zelealem (2009) describes lexical development in Tigrinya. Accordingly, various lexical elaboration mechanisms in the language are identified. The result of the study depicted that compounding, semantic extension, borrowing, loan translation, and rarely blending and acronymy are identified as the principal lexical elaboration mechanisms of the language. Despite an existence of genuine words of Tigrigna, unnecessary borrowings are employed to expand its lexicon and their replacement by native words is also suggested in the study. Zelealem (2009) also identifies that the participation of various stakeholders (concerned bodies) like educators, linguists, and elderly speakers of the language (rural and urban) is not to the extent needed. As a result, he suggested that participation of various individuals from various kinds of expertise is vital for the effective achievement of lexical enrichment of the language in the period to come.

Zelealem (2012) examines the typological comparisons of the language policy of the Imperial, the Derg and the EPRDF governments. The result of the study indicates that the three regimes followed different approach to maintain the unity of the country despite the fact that they have similar interest to sustain national integration. The result of the study shows that the unity and strength of the nation was assumed to be maintained through the solely use of one language (Amharic) during the Imperial government. Language diversity was also considered as a threat to

the strength of the country and diversity was not appreciated in general. As a result, an assimilationist language policy was encouraged. During the Derg and EPRDF regims, however, language diversity is considered as a norm. Both these two regimes have included the diversity in their respective constitution. The right of ethno-linguistic groups to use their languages and promote their cultural heritages is language diversity related issues that is clearly stated in the respective constitution. This is seen to be practical as many languages are used in the media and education. Various languages, like Oromo, Tigrinya and Somali were used as a medium of instruction particularly in primary schools. Due to such practical language development activities, EPRDF is better in its language diversity policy than other regimes.

Bekale (2012) describes the current day language planning (LP) of Ethiopia. The target of the description is to identify LP strengths and weaknesses as well as the challenges and problems of the then period of the country. The result of the study indicates that LP of EPRDF shows better success than the Imperial and Derg regimes period. As per the finding of the study, two serious problems are identified as obstacles to the effort of LP of the current period. Firstly, there is lack of systematic, theoretically sound, coherent, consistent and sustainable LP. Secondly, language planners have lack of skill and knowledge which has resulted in planning the language in an unprofessional approach (trial and error methods). Bekale (2012) proposes two solutions (i.e., long and short term) to address the above problems. Firstly, providing language planners with an easy and friendly fact of LP model; based on the descriptive framework is suggested as the short-term solutions. This is due to the fact that language planners can benefit from the main fact of LP model and use it as a rule and instructions to accomplish their task. Bekele (2012) also claims the availability of a guideline which can satisfy the growing needs of a multilingual developing country like Ethiopia. Hence, he outlined a guideline on the basis of a descriptive framework of language planning that he calls the Modified Descriptive Framework of LP. Secondly, the long - term solutions like provision of professional development program through training in the area of LP is also recommended so to solve the stated problems.

Shimelis (2016) describes the salient features of Hadiyya orthography and discusses aspects that seem to require improvements. To investigate this, a short Hadiyya paragraph is dictated to and is written by randomly selected elementary, high school and first-year university students. The

result of the study shows that the students had problem of distinguishing phonological quantity, that is, between simplex and geminate consonants as well as between short and long vowels. As a result, representing geminates and long vowels as simple segments occurred frequently. The findings of the study also show that some aspects of orthographic inconvenience which require adjustments are encountered and solutions are suggested from the perspectives of regularity and economy. The study also identified the constraint in relation to the transferability of incompatible knowledge of Hadiyya to English. As per the finding of study, the performance of female students, especially at high school and university levels, is found to be lower than that of the male students and the study suggests that this needs to be further investigated. He finally, recommended that the strategy has to be devised to solve the problems related to the orthography and its use a standard dictionary has to be compiled.

Samuel (2017) has conducted a PhD research to make a contrastive analysis of lexical standardization in Amharic and Hadiyya. The research has shown that the overrepresentation of sounds, i.e. /h/, /s/, /s'/ and vowels, became a problem of standardization of Amharic spelling. The result of the study also shows that the spelling variation in Amharic is due to unnecessary retention of symbols during its development of writing system from the Ge'ez script. In Amharic, spelling variation is attested in compounds and loan words. The author also states that the violation of the principle of economy and transfer is attested in the Roman-based Hadiyya orthography. The research has also shown that the gemination of digraphs reduces economy while the assignment of different phonetic values to some of the Hadiyya and English sounds negatively affects the transfer of reading and writing skills. The result of the study depicts that the graphemes <zh> /ʒ/ and <ny> /ɲ/ are overrepresentation in Hadiyya. As per the finding of the study, language internal and external resources were extensively used to elaborate the language. The result of the study has also shown that semantic transfer and compounding strategies are predominantly used for lexical elaboration in both languages whereas abbreviation and blending are less common.

Hadiyya extensively employed semantic transfer method more active than Amharic. On the other hand, borrowing and loan translations are identified as the most productive language external means of lexical elaboration in both Amharic and Hadiyya. Both languages borrowed words

mainly with modifications; but unmodified borrowing was rarely found in Amharic. To the contrary, this is not attested in Hadiyya. The result of the study also shows that foreign languages, such as English, Arabic, French and Italian are the main source of borrowing than Ethiopian languages. Amharic exploited the genetically related language Ge'ez to enrich its vocabulary, whereas Hadiyya does not exploit related Cushitic languages. As per the finding of the study most of the European loanwords entered Hadiyya via Amharic. Regarding the main source language of loan translation, English and Amharic are used as source languages for Hadiyya whereas English is used as a source language of loan translations for Amharic.

In her PhD study, Almaz (2018) investigate the social challenges faced in the standardization process of the Gamo language. The finding of the study shows that the norm selection technique followed in the standardization of Gamo was a monocentric one. The study reveals that “the standard form of Gamo” was mainly based on the Dache dialect, which is one of the dialect variants of the language spoken by the majority of the population. As the finding of the study shows, the majority of the respondents had positive attitude towards Dache to serve as a standard Gamo. The study considers that the adoption of a polycentric multidialectal approach in the establishment of standard Gamo is one of the basic factors to achieve successful language standardization. The study also identified the challenges of standardization of Gamo. Accordingly, the degree of dialectal variation, procedures followed to standardize the language, and attitudes held towards using Gamo in wider public settings were considered to be the major challenges of standardization in Gamo.

2.2.3 Review of studies on standardization of Oromo

This study found the inclusion of an overview of history of the standardization committee of important. However, it is difficult to find literatures about the history of the committee even in the materials published by the committee. Therefore, a very short overview of history of the standardization committee of Oromo is given. It was in 1994 that the standardization committee of Oromo (Koree Waaltina Afaan Oromoo) established by the then Oromia Culture and Information Bureau (Mekonin 2002:74). The main task of this committee is to standardize the language as the language has several dialects and variation that may lead into mutually

incomprehensible languages. The then members of the committee were selected from the Oromia Culture and Information, Education and Judiciary Bureaus as well as from the regional government offices and they were few (did not exceed 7). Less effort was given to fulfill the human and non human resource needs of the task. That is why the committee codified/coined only 1500 words within the years 1995-2000 (see Wiirtuu, 1995-1999, No.1-5). The five year output of the committee is almost few as compared to other countries committees' or academies' output. For instance, the academy of South Africa as stated by Webb (1995) produced 10 different bilingual glossaries (English-Afrikaans) within only 6 years (1994-1999). Similarly, in the effort to standardize the Hindi language, the language authority produced 200,000 terms related to science, humanities and social science out of which 125,000 were published as it is stated by Gupta (1969:578), within only 11 years (1967-1968). On the other hand, the codified words of Oromo by the standardization committee of Oromo for the last 27 years (1994-2021) do not exceed 10,000 words. The standardization committee of Oromo published 11 *Wiirtuu* (Journals) during these 27 years. The contents of these *Wiirtuu* (journal) include literature and language standardization. Here, we can note that the committee seems to lose its main task i.e., working on normalization process of the language. What has been done is almost none as compared to what has to be normalized. There are several reasons for that. First, the numbers of members of the committee is very few as compared to tasks to be done. There are only 12 members all of which are masters and degree holders. There must be doctors and professors and their number has to be as many as possible (at least 40 or 50) to accomplish this challenging task of normalizing the language. Second, the budget allowed for task is not enough. Third, the task of the standardization of the language is being done at the level of committee for the last 27 years (1994 up to today (2021)). It has to be elevated to the level of language academy. There are also challenges of implementation (*Wiirtuu* volume 10 2005:45). What is codified is not seen to be implemented as stakeholders are not working hand in hand. All in all, though the committee played its own role, the standardization process of the language is not to the extent needed. Hence, this study recommends that the concerned body (Oromia Culture and Tourism Bureau) has to work hard to resolve all those shortcomings so as to standardize the language.

Temesgen (1993) examines word formation in Oromo. He describes the derivational and compounding methods of word formation in Oromo. He identified different types of nominal,

verbal, and adjectival those are formed by the process of derivation. He further states the existence of varieties of affixes which are used to derive nominals. He also states that the distributions of nominalizing suffixes are not predictable and that the derivation of verbs is complex. Most adjectives of the language are basic and the few derived ones are formed with the analogy of these basic ones. Compounding is another word-formation process treated under the study. The result of the study shows that nominals, adjectivals and adpositionals are among word formation methods that are formed through the process of compounding. Finally, he concluded that determining the head of compounds in general is difficult as it needs more thorough investigation.

Baye (1994) examines lexical development in Oromo and mentioned what was being done in the standardization and lexical modernization of the language. He also puts an idea forward ways that could be employed in order to coin the needed technical terms. The result of the study suggests that borrowing, loan-translation, compounding and derivation are elaboration mechanisms used for lexical development of the language. He further suggests that it seems important to see from the available varieties of the language before looking into loanwords of foreign origin even though borrowing is natural to enrich the language.

In his MA study, Mekonnen (2002) examines lexical standardization in Oromo based on Huagen's (1966 & 1969) four stages of normalization processes and propose all varieties as a base for the standardization of Oromo. Regarding the methods of lexical elaboration, he makes the conclusion that blending, semantic extension, compounding, derivation and borrowing can be used in developing the lexical adequacy of Oromo in expressing scientific, technological and other concepts. Finally, to implement the proposed standard form, Mekonnen (2002) explains that we can expect much from governmental institutions like the education bureau, the mass media, a language academy/committee, and non-governmental organization such as that of the missionaries, individuals and the speech community itself. Dictionary compilation is one to language standardization. The study also shows that efforts have been made by various authors to harmonize the language through the inclusion of dialects of the language in their dictionary. As per the finding of the study, the pre 1991 dictionaries were one dialect based to the contrary to the post 1991. The selection of a base dialect or base dialects is carried out on the basis of

sociolinguistic factors. The factors are: written documents, the mass media, current standardizing tendency and the attitude of the speakers. The study showed that all dialects should be considered in standardizing the language.

Fikadu (2010) examines spelling and grammar errors of Oromo modules prepared in Haramaya University. The finding of the study depicts the existence of spelling errors such as omission, addition, analogy, substitution, transposition, spacing, inconsistency and unclassifiable error types. The research identifies grammar errors such as incorrect use of connectives, case markers, parallel constructions, agreements, prepositions and tenses. The findings also reveal that these errors affect students' understanding in different ways. The study also shows that carelessness and lack of knowledge which results from absence of practice, good Oromo academic background and insufficient experience were found to be the causes of the errors in the modules.

Amanuel & Samuel (2012) examine homonymy as a barrier to mutual intelligibility among speakers of varieties of Oromo. The finding of the study shows that lexical, phonological and morphophonemic variations in Oromo can give rise to homonymous and homonymy-like lexical items that pose difficulty to mutual intelligibility among speakers of different varieties of Oromo. The authors further claim that the phonological variations can emanate due to three linguistic phenomena i.e., tonal difference, h-dropping or using /h/ and /ʔ/ interchangeably at the beginning of words, and avoiding *dh-* or replacing the phoneme /dh/ by the glottal stop /ʔ/ in some dialects. These three linguistic phenomena will create phonological variation that give rise to ambiguous expressions. Amanuel & Samuel (2012) further state that in the morphophonemic of some variety, assimilation is the common process that results in the formation of homonymous expressions and the morpheme that plays a significant role in the process is first person plural marker *-n-*. As to the authors, the convention in written Oromo which allows speakers to write almost exactly as they pronounce it and lack of standard Oromo are the main reasons for the occurrence of homonymy-like lexical items that contribute to misunderstandings and confusions among Oromo people from differing dialect areas.

In his MA study, Teshome (2014) examined lexical use of Oromo in the government mass media. The study identified that due to the existence of varieties and lack of a standard form of

Oromo, there is no uniform use of the language in government mass media. The finding of the study also shows that due to the fact that there is no frame of reference that the editors use to edit words used by reporters, journalists use their own dialect which may not be fully understood among speakers of other dialects. The study finally, recommended that there should be a style guide to which editors and journalists refer to so as to apply standard Oromo in the government mass media as in the case of, for instance, BBC, Reuters, and The guardian.

Tamam (2015) described issues in the Standardization of the Oromo Orthography. In his major research paper, Tamam discusses about the name and graphematic representation of the glottal stop, germination of digraphs, tone, punctuation marks and word segmentation issues in Oromo. Regarding the phoneme inventory; he suggested that the glottal stop /ʔ/ currently represented by the apostrophe <'> and named differently by different authors be named hamza following Moreno (1964) and assigned the first phoneme in the alphabet. He also suggested to avoid writing it word initially just like the suggestion given by the present study. Regarding the gemination and order of the digraphs; other than voiceless palatal stop /c/ and palatal nasal /ɲ/ which are geminate word internally, Tamam recommends first letters of the digraph to be doubled for germination. Regarding the punctuation marks; he suggested to fully adopt the English rules since English is currently the language of higher education in the country. As for the tone; he states that even though the debate as to the tonal nature of the language is not resolved, it is time to start including it in the curriculum of the language teaching and mark it in dictionaries at least when there are minimal pairs. Regarding word segmentation; he suggested that disambiguation rather than 'we write as we speak' motto should be followed. Tamam finally recommended that the preference of writing certain morphemes, mainly clitics, disjunctively with few suggested to be conjoined with hyphen,

Teferi (2015) discusses the development of Oromo writing systems in which the language remained officially unwritten with officially accepted script until the early 1990s. The study begins by reviewing the early history of Oromo writing and discusses the Ethiopian language policies, analyzing materials written in various scripts and by certain writers starting from the 19th century. He states that the development and use of languages for office, education, religion, etc. purposes have been a major political issue in many developing multilingual countries. Teferi

(2015) also states that countries, such as Ethiopia (a multilingual African state) had, for a long time, preferred a policy of one language and one script in the belief that this would help the assimilation of various ethnic groups create a homogenous population with one language and culture. The author argues that the Roman based alphabet is the best option for Oromo writing as it is relatively easy to adapt for representing the phonological aspects of Oromo. The study argues that Oromo has thrived during the past 23 years to implement a Roman based alphabetic script. It is also stated in the study that there have been and continue to be, however, internal and external challenges confronting the development of the Oromo writing system which need to be carefully considered and addressed by stakeholders, primarily by the Oromo people and the Ethiopian government, in order for the Oromo to establish itself as a fully codified language in the modern nation-state. Teferi (2015) argues that while the Roman alphabet is the preferred script for Oromo writing, there still remain formidable challenges to be faced in making it a viable system. One of these challenges is the issue of the standardization of the language. A standardization of Oromo is needed for political as much as for linguistic reasons. The author further recommended that detailed studies on standardization of Oromo is needed.

Temam (2018) describes copula in Oromo which he claims the unsettled issue in the grammar of the language. The author argues that there is little agreement on its existence or what represents that. The case of copula that attracts attention is the one in the simple present tense. To him, many authors consider particles like *dha*, *ti* and *tu* as well as null (o) and also *miti* 'not' in some cases, while few consider the low tone on the final vowel on the prepausal nominals. The finding of the study shows that a copula or a predicative case is represented by neither the particles nor by the low tone; but by the short vowel at the end of a predicative nominal. He also claims that some of the components or some of the particles considered as copulas are only epenthetic added to introduce the actual representation of copula; which is a short vowel that cannot directly be added to an existing vowel at the end of the nominal.

The present study has a similarity with Temesgen (1993), Baye (1994), Mekonen (2002), Fikadu (2010), Teshome (2014), Teferi (2015) and Amanuel & Samuel (2012) in that they all deal with issues that are related to language standardization. But, there is clear difference regarding the focus of study. Temesgen (1993) focuses on the word formation of Oromo. Baye's (1994) study

on the lexical modernization of Oromo whereas Mekonen (2002) deals with how lexical standardization in Oromo can takeplace. Mekonen's (2002) study is all about the propositions to establish criteria which help to choose standard words out of different dialects practical methods of coining new words, and its implementation. The focus of Fikadu's (2010) study is on the failure of the module compilers to conform to the canonical rules of Oromo during the preparation of modules. Target of Teshome's (2014) study is on lexical use of Oromo in the government mass media and it is all about the lexical variation in media whereas Amanuel & Samuel's (2012) work is far from the rest of the studies in that it examine homonymy as a barrier to mutual intelligibility among speakers of varieties of Oromo. Though, Tamam's (2015) study is directly related to the present study, there is clear difference as this study is on orthographic and lexical standardization of Oromo based on text corpus. Hence, here, the present study has a clear departure from all the rest of the studies stated under this study. This study focuses on the orthographic and lexical standardization of Oromo based on a text corpus from the standardization committee of Oromo and from the current primary and secondary textbooks prepared in Oromo as well as based on data collected through FGD and interview.

CHAPTER THREE

ORTHOGRAPHY DEVELOPMENT OF OROMO AND ITS STANDARDIZATION

3.0 Introduction

An orthography is a specific set of symbols and the rules used to combine those symbols into words and sentences for a specific language (Decker 2014:25). Orthography also includes the rules that map the graphemes to sounds and the patterns by which the graphemes can be combined to form words (Decker 2014:25). A standardized orthography includes all of the socially accepted spellings for all of the words in the language. In addition to the syntactic rules of word combination in a language, there are also orthographic rules that become normative for writing texts (Decker 2014:25). These orthographic rules include the conventions for using paralyphs, the non-alphabetic symbols. The critical issues that are related to historical overviews of the development of Oromo to a written language, arguments over script choice and the development of Roman-based script and evaluation of the current Oromo orthography on the bases of the principles of orthography development are presented and discussed in the following sections.

3.1. Oromo phonology

Before engaging into the standardization issues in Oromo, discussing the phonemic inventory and the phonology of the language is vital. The phonemic inventory and phonology of the language will be discussed briefly in the section. Oromo consonant phonemes are shown in the following table.

Table 4. Consonants of Oromo

		Bilabial	Labio-dental	Alveolar	Palatal	Velar	glottal
Stop	Voiceless	(p)		t	tʃ	k	ʔ
	Voiced	b		d	dʒ	g	
	Ejective vɪ	p'		t'	tʃ'	k'	
	Implosive vd			d̥			
Nasal	Voiced	m		n	ɲ		
Trill	Voiced			r			
Lateral	Voiced			l			
Fricative	Voiceless		f	s	ʃ		h
	Voiced		(v)	(z)			
Glides	Voiced	w			j		

Source: (Gragg 1976:174)

The above chart of phonemes of Oromo is taken from Gragg's (1976: 174) phonemic inventory of Oromo. The phonemes [ɲ], [ʃ], [p'], [tʃ'], and [d̥], are written in the Oromo orthography using digraphs as <ny>, <sh>, <ph>, <ch>, and <dh> respectively. Phonemes [k'], [p'], [t'], [tʃ'], are ejectives with an egressive glottalic initiation; there is upward movement of vocal tract during the articulation of the phonemes. In Oromo, the alveolar implosive phoneme d̥ is the only sound articulated by glottalic ingressive suction initiation with vocalic articulation. The Oromo has two glottal segments, the fricative /h/ and the stop /ʔ/, both of which occur only non-geminate. This feature makes Oromo glottal segments different from the other consonants which generally occur both in single and geminate forms (Teferi 2015:37). The standardization committee of Oromo codified the consonants of Oromo like exactly as it is depicted in the above consonant inventory; but added one more consonant, the alveolar ejective, /s'/ as loan phoneme and represent it using grapheme 'ts' making the number of Oromo consonants twenty eight (Qajeelcha 1998:14). The grapheme of this loan phoneme, however, not included in the alphabetic order of the language. This could be due to the fact that there are only few loanwords in which this phoneme occurred.

As stated by scholars like Kebede (2009:26) & Shimelis (2014:4), there are twenty-four native consonant phonemes in Oromo. The standardization committee of Oromo considers /s/, /z/, /v/ and /p/ loan phonemes that enter the language as a result of lexical borrowing. The last three loan phonemes are represented by <z>, <v> and <p> and are included in the alphabetic order of the language. Oromo has two mechanisms of dealing with loan phonemes, i.e. either replacing a loan phoneme with the nearest native phoneme or adapt it as it is as in, for instance, Amharic loan phoneme *tsabala*→*xabala* 'Holy water' and the English loan phoneme /v/ *virus*→*vaayirasii* 'virus' respectively. The English loan phoneme /ʒ/, which is neither yet declared as a loan phoneme nor codified to be represented by a grapheme by the standardization committee of Oromo. This loan phoneme occurs in primary and secondary textbooks of Oromo with both replacement of native nearest phoneme *j /dʒ/* and with adaptation via representing it with <zy>, <zn> and <zh> as in, for instance, *televijinii*, *televizyinii*, *televiznirii* and *televizhinii* 'television' respectively (EBO Environmental science grade 5: v and 165) (for loan phonemes and their phonological processes in Oromo in general, see section 4.2.1.3 and 5.2.1.3.1.1.2). There is variation in representing this loan phoneme while writing loanwords that contain the phoneme and an appropriate grapheme is not yet designed by the standardization committee of Oromo.

Regarding vowels, Oromo has five short vowels with corresponding five phonemic long vowels. Length is represented by doubling the short vowels.

Table 5. Oromo Vowels

	Front	central	Back
High	i		u
Mid	e		o
Low		a	

Source (Qejeelcha 1998:8)

Gemination of consonant and vowel length in Oromo is marked by doubling the respective consonants and vowels. Consonant gemination and vowel length in Oromo are phonemic. Oromo permits a maximum sequence of two different consonants within a word, but neither at the beginning nor at the end of a word (Teferi 2015:50; Wako 1981:34). The epenthetic vowel [i]

settles such an occurrence. The following tables show examples the consonant gemination and vowel length in Oromo

Table 6. Consonant gemination in Oromo

Status of the phoneme	Oromo	gloss
Non geminated	<i>hatuu</i>	'to steal'
Geminated	<i>hattuu</i>	'thief'
Non geminated	<i>sodaa</i>	'fear'
Geminated	<i>soddaa</i>	' mother-/father-in-law '

Source of data (Qejeelcha 1998:24)

Table 7. Vowel length in Oromo

Status of vowel phoneme	Oromo	Gloss
short	<i>mala</i>	'method/technique'
long	<i>maala</i>	'dewlap of ox'
short	<i>lafa</i>	'land'
long	<i>laafaa</i>	'soft'
short	<i>boru</i>	'tomorrow'
long	<i>booruu</i>	'unclear water'
short	<i>gale</i>	'I/he entered'
long	<i>gaalee</i>	'/phrase/shield'

Source of data (SCO Wiirtuu volume 10 2005:15; Qejeelcha 1998:24)

3.2. Syllable structure and vowel harmony of Oromo

There are four types of syllable structures in Oromo. These are CV, CVC, CVV and CVVC. Onset position is always occupied because neither word nor syllable begins with a vowel (Wako 1981:37). Griefenow-Mewis (2001:21) also states four types of Oromo syllable structure namely CV, CVC, CVV and CVVC. In the following transcription the dot (.) between two syllables represents the syllable boundary and the bold syllables are the example syllables under their respective syllable types. The following examples are taken from (Qajeelcha 1998; Wiirtuu volume 5 1992).

(1)	Oromo		gloss
1. CV	<i>nama</i>	na.ma	'person'
	<i>harra</i>	har.ra	'today'
	<i>bakka</i>	bak.ka	'place'
2. CVV	<i>haaraa</i>	haa.raa	'new'
	<i>eegee</i>	?ee.gee	'tail'
	<i>caasaa</i>	tf'aa.saa	'structure'
3. CVC	<i>kun</i>	kun	'this'
	<i>karra</i>	kar.ra	'door'
	<i>harma</i>	har.ma	'breast'
4. CVVC	<i>bishaanbi.</i>	faan	'water'

As can be seen from (1), a syllable in Oromo often formed from onset, nucleus and coda. Oromo disallows empty onset since the phoneme /ʔ/ considered as the first member preceding a vowel as in *?ee.gee* 'tail'. Regarding the co-occurrence of consonant phonemes, Oromo allows only two identical and different consonant clusters only intervocalically in a word as in *harra* 'today' and *harma* 'breast' respectively. The occurrence of a cluster both word-initially and word-finally is strictly forbidden in Oromo and will be conformed by placing the epenthetic vowel *i* before the cluster in word-initial position and after the cluster in wordfinal position (Shimelis 2014:6). Oromo words can end with a single consonant or a vowel (s) as in *kun* 'this' and *nama* 'person' respectively. Vowel length Oromo can occur word medially and word finally as in *caasaa*

'structure' and a single vowel (short vowel) can also occur word medially and finally as in, for instance, *fala* 'solution' respectively. But vowel length is disallowed word initially.

Apart from the vowel length, phonotactic, syllable structure of Oromo, a short overview of vowel harmony is essential for the present study as there are phonological adaptations during borrowing in the language as we shall see under external means of lexical elaborations in sections 4.2.1 and 5.2.1.3.1 of the fourth and fifth chapters respectively. Generally, vowel harmony is a phonological process that stipulates that vowels within a particular phonological, lexical and sometimes syntactic domain (word/phrase) must share specific features (Akanlig-Pare & Asante 2016:1). In other words, vowel harmony is patterns where in vowels assimilate with respect to some feature and the assimilation could be adjacent or distant. Features shared by vowels in harmony may include [+/-ATR (advanced tongue root)], [+/-high], [+/-back], and [+/-round] (Akanlig-Pare & Asante 2016:1). Advanced tongue root (ATR) and retracted tongue root (RTR), are contrasting states of the root of the tongue during the pronunciation of vowels in some languages (Akanlig-Pare & Asante 2016:1). These authors also state that the [+ATR] vowels are produced with an advancement of the tongue root while the [-ATR] vowels are articulated with retracted tongue root. Vowel harmony is seen in Oromo during the integration of foreign lexical items into the language as in, for instance, the Amharic loanword in Oromo *sik^wuar*→*sukkaara*.

3.3 Development of Oromo to a written language

This section of the chapter discusses the general overview of early Oromo writing. This section analyzes the early individual authors' attempt, arguments over script choice and the development of Roman-based script to write Oromo.

3.3.1 Early attempts

It is widely accepted that Oromo have not been written until the early 1840's and the first available evidence of written Oromo is a book on the elements of the language (vocabulary of Oromo) was written by Krapf in 1840 (Krapf 1840:4). The author further states that it was the European missionaries, scholars and diplomats, who were responsible for making pioneering

efforts to translate Oromo into written from the 19th century and the Europeans' interest in the Oromo was mainly to spread the Word of God amongst the Oromo people as well as their desire to teach the language to their fellow Europeans.

Oromo was written using Arabic, Ge'ez and Roman based script before 1991. For instance, Krapf (1842) is one of the authors who had tried to write Oromo using both Ge'ez and Roman-based script; but he did not clearly state why he had selected the Roman based alphabet to write his 'Vocabulary of the [Oromo] language', and the Ge'ez script to write the 'Gospel according to St. John'. Other authors like Hudson & Walker (1922) also used the Roman based alphabet and Ge'ez script like the previous author. Contrary to Krapf (1842), Hudson & Walker (1922) state why they used both scripts.

It was at one time intended to print all the sentences in the Ge'ez or Amharic characters as well as in Roman, but the idea was abandoned as impractical so far as the more grammatical part of the work was concerned.

(Hudson & Walker 1922: 9)

The authors further explained that the use of the Ge'ez scripts can help Oromo people to learn Oromo as there were a number of Oromo people that had already read and write using the Ge'ez script. On the other hand, Tutschek (1845:1) used the Roman based script to write Oromo rather than Ge'ez script and he also put his justification for not using the Ge'ez script to write Oromo. The author states that if the [Oromo] has a written language, which is rendered possible by recent accounts, it must have a peculiar alphabet or syllabarium¹ that express all the sounds that are found in it with exactness. To the author, the [Oromo] cannot be written with the Amharic syllabarium as it fails to represent many of the sounds to be met with in [Oromo].Tutschek (1845:1), finally, concluded that the adaptation of Roman characters to [Oromo] sounds, as far as it is possible, is the only way of making the language accessible (Tutschek 1845:1).

Tutschek's (1845:1) view is clear; Oromo cannot be written in Ge'ez as it cannot express all the sounds in Oromo exactly and the solution is writing Oromo using the Roman script. Onesimosi was one of the authors who have played a role in writing Oromo using the Ge'ez script. But he

¹syllabarium, Ethiopic and Sabeian script are other names for Ge'ez script

did not justify why he had selected the Ge'ez script to translate various materials in Oromo. The reason for his choice of the Ge'ez script might be due to the fact that the then Oromo peoples were exposed to the Ge'ez script or he might be inspired by (Krapf 1842) who translated faith books into Oromo using the Ge'ez script. Onesimosi translated different materials; both faith and secular books (Onesimosi 1899). For instance, a faith book, the Bible, into Oromo using the Ge'ez script and most of them were translated between 1885 and 1898 (Mekuria 1994:94). A dictionary known as 'The [Oromo] Spelling Book' was also one of the books among the secular books Onesimosi wrote (Onesimosi 1894).

Sheik Bakri Sapolo was another notable author who wrote Oromo through developing an Arabic script in order to write Oromo. As to Hayward and Mohammed (1981), Sheik Bakri's works deal with secular as well as religious subjects, were written mainly in the Arabic, and the Ge'ez script. Due to various reasons, although a few of them have been typed and mimeographed, most of his writings were circulated in manuscript form only among his disciples, and so remained unknown to the wider public. This is because the works are written in Arabic which is not read and understood by the majority of the people.

Sheik Bakri Sapolo devised an indigenous and original script as part of his attempt to overcome problems of orthography in writing Oromo as he previously wrote Oromo in Arabic (Feyisa (1996:22)). He devised a script which was different in form, but followed the symbol-sounds forming patterns of the Geez system. Having developed the alphabet, Sheik Sapalo taught it to all his students and then, people began to exchange letters in the new alphabet (Wondimu 2016:12).

Generally, it can be concluded that Sheik Bakri Sapolo made an effort and motivation to develop literacy among the Oromo (particularly in Harar). Through understanding that the existing Arabic and Ge'ez scripts did not fit well with the peculiar features of the sound (phonology) of Oromo, Sheik Bakri developed a unique writing system for Oromo in the early 1950s. For instance, the main deficiency of the Arabic script is the problem of marking vowel. The Ge'ez script does not differentiate gemination of consonant, vowel length and glottal stops of

Oromo. Hence, the author's effort can be taken as an innovation that surpasses the context of an introduction of Qubee of Oromo afterwards.

During the 1960's till 1990's various authors and organized bodies used both Ge'ez and Roman based scripts to write Oromo. For instance, OLF (Oromo Liberation Front) published using Roman script and distributed an Oromo textbook called '*Barmoota Afaan Oromoo, Barreeffama Lammaffaa*,' in 1980 for teaching Oromo in the 1980's (Mekuria 1993:15). This shows that Oromo was written using Roman script some eleven years back before the Roman script was officially adopted in the country, in 1991, as a script on which the language writing system was to be based.

3.3.2 Arguments over choice of script and the development of Roman-based script

3.3.2.1 Arguments over the choice of script

Before directly going to discuss the existing arguments regarding the choice of script, articulating the general principles of orthography is vital. The design of orthography is connected with various issues like linguistic, political (the internal dispute and the influence of the outside world). There are general principles that have to be followed while designing orthography in order to write a language. Linguistic, socio-political (acceptable to all stakeholders), pedagogical (teachable) and practicality i.e.technical aspects or easy to reproduce (Cahill & Karan 2008:3)

The linguistic principle is among the fundamental principles which need due attention during the design of orthography to write a language.As a general principle, contrasting sounds in a particular language have to be represented distinct symbols (Cahill & Karan 2008:3). For instance, in Oromo distinct phonemes like /l/ and / n/ need independent graphemes <l> and <n> respectively as they contrast as in *lama* and *nama*. Apart from the design of a script to write a language, word division related issues are also one of the linguistic principles to be considered during the design of orthography.Word division tasks include whether two or more morphemes to be written together (attached) or should be spelled separately. It has to be also stated clearly as to how compounds, clitics, pronouns, affixes, supra-segmental such as tone, nasality, and vowel harmony should be written (Cahill & Karan 2008:6). Minimization of potential ambiguity of words in context and simplicity of orthography are also the elements of good orthography.

Simplicity of orthography is related to its readability, (teachability and learnability) of orthography (Powlison 1968:80).

The pedagogical issue of orthography is another pillar that has to be considered during the development of orthography. Maximum ease of learning, maximum ease of transfer to other languages (Malone 2004:8) is the pedagogic consideration of orthography design. The inherent difficulty or ease of orthography is related to its learnability. Orthography has to be consistent and as simple as possible in order to attain maximum ease of learning (easy to read and to write) (Tench 2007:2). The mismatch between the written and the spoken language will definitely create difficulty in learning to write and read. If there are orthographic underrepresentation (results when an orthography uses fewer graphemes than there are phonemes), overrepresentation (using more graphemes than there are phoneme), inconsistencies in sound-symbol correspondences, a silent letters, and loan words retaining unadapted spellings, then, there will be deep orthography (Tench 2007:2).

The maximum ease of transferability is another element of pedagogical issues of Orthography. Obviously, human being wants to learn another language in addition to his/her own language for several reasons such as, for economic, political and social (to make oneself a part of another societies life) benefits. Hence, an orthography of a language has to be consciously designed the symbols and all other conventions one language so as to help learners and teachers to transfer skills to other languages (Cahill & Karan 2008:8).

Orthographic design should also go beyond the attainment of linguistic and pedagogical requirements. Orthography has to be socially acceptable. The user of language demands orthography of their language not to forget their history, culture, religion and identity. As to Cahill & Karan (2008:9), orthography could be acceptable to some speakers of a dialect while being unacceptable to others and hence, finding the acceptable via considering the size, location and prestige of the dialect are vital. Thus, like other factors we have seen, the social acceptability of orthography is very important (if not more important than the rest of the factors stated above) factor in designing an orthography. Hence, the development of orthography has to make sure that the users of that orthography have a positive attitude towards the designed orthography, pride, a

reasonable interest to document and use their written form of language on the day to day bases (Cahill & Karan (2008:10).

Practical production factor is also another important issue to be considered during the designation of orthography. Previously, the choice of orthography depends on its suitability with the existing technology (Cahill & Karan 2008:9). Symbol options depend on a typewriter or local printer capabilities. There was a challenge of adapting Roman alphabet and some symbols to the then technology. Hence, due attention has to be given to the adaptability of the orthography to the reproductive factors. Orthography of a language has to be designed by taking into consideration that graphemes conform to the already available technological reproductive devices such as standard keyboards.

All in all, each factor of orthography development has to be given attention while designing orthography. The task of an orthography developer is to evaluate the above discussed factors and then try to find the middle place (balance) among the stated elements of orthography design as some these factors seem conflicting. For instance, the users of a language might want their orthography to be distinct from a dominant culture and at the same time want their orthography to have a feature of maximum transfer to other languages. Hence, orthography has to be accurate (important sound differences are recognized and written in a distinctive way), consistency (the same symbol always represents the same sound), convenience (any special symbols used should be easy to type and keyboard on a type writer or computer) and conformity follow the writing system of the language of wider communication in the area to the extent possible (Barnwell 2004:5).

Coming back to the argument on the choice of script to write Oromo, there are a number of scholars who argue for and against both Ge'ez and Roman based script. The arguments given by various scholars for and against of the two scripts for writing Oromo can be categorized into four major points namely: linguistic, pedagogical, practical and attitudinal.

It is obvious that Roman-based script and Ge'ez script were among the scripts that were suggested to be the competing scripts for writing Oromo during the meeting that was held to decide the official

script of Oromo. Baye (1992:5) was one of the earliest linguists who argued that the Ge'ez script can be used to write Oromo with improvement and there is no linguistic reason to resort to the Latin script. It is true that Ge'ez script is weak in marking gemination and vowel length without making adaptation and hence, alternative meaning can be created (words are subject to different interpretations). The scholar, however, further argues that the problem of alternative meaning that might arise due to the unmarking of consonant gemination and vowel length can be known from the context in which the word is said. He also adds that such problem of unmarkedness of gemination is also seen in Amharic and English too. For instance, the Amharic ገና can be geminated as in 'gəṇna' (christmass) or not 'gəna' (not yet). Here, the author noted that whether it has to be geminated or not and short or long can be known from the context in which the word is said. Baye (1992:4) further explains that the Latin script is also not immune the unmarkedness. He has given an example from one of the language (English) written using Latin script words like "export", "import" and so on are read (understood) either as verb(s) or as noun(s). This can tell us that there are several such instances in which the meaning and the category of the words are known from the context they occur and not from the script.

Apart from the remedy of reading (understanding) the consonant gemination and vowel length via looking into the context a word occurred, dot (.) is also suggested to mark on the top or beside of the Ge'ez character to be geminated or the vowel to be lengthened (Baye 1992:5; Launhardt 1973:221; Teferi 2015:133). Here, Baye (1992:5) mentions that though putting dot on the top of a Ge'ez character to be geminated or lengthened, such remedy does not look feasible because of the belief that it will require additional symbols on current computer. This is actually not a linguistics difficulty; but technological and should be improved by technologist via developing a software to efficiently write Oromo with Ge'ez script just like it has been done for Amharic (Baye 1992:5). He argued that though the computer is based on the Latin script, the software itself converts each typed Latin script to that of the Geez. Baye (1992:5) also mentioned another method of solving the problem of marking geminated and long sounds while writing Oromo with Ge'ez script. To him, writing the "sads" (6th order) consonant of the geminated sound and vowel length as in, for instance, "ቦርእኑ" (written document) can represent that consonant 'r' is geminated and the vowel is long. But, here, the present researcher did not find an argument by

any scholar as to the Oromo implosive sound /d/ and the glottal stop /ʔ/ be represented if the Ge'ez script is to be adapted to write Oromo and this remain a problem.

On the other hand, there are also scholars who argue in favor of Roman-based script to write Oromo. The criticism against the Ge'ez script is that it has limitations to adapt for writing Oromo such as disability of marking gemination for the consonants and its failure to show the length of vowels as well as its inadequacy to represent some Oromo sounds (the implosive and the glottal stop) (Wodimu 2019:18; Teferi 2015:134; Tilahun 1993:37 & Gadaa 1988:16). The representation of consonant gemination, vowel length and the implosive and glottal stop is, as can be seen from various linguistic scholars, could be arguable. But, there is linguistic misconception from those in favor of Roman-based script to write Oromo. The linguistic economy is one of the issues to be considered in orthography development for a language. From this point, Feyisa (1996:26) argues that Roman-based script to be every economical than the Ge'ez script. This argument is an absolute misconception as Roman based script is not economical as compared to the Ge'ez script when it is judged from the number of graphemes required to write a word. For instance, the word 'house' is spelled in Oromo as *mana* and ሙን in Ge'ez script. Here, one can easily note that the Roman based script has to use more (four) characters to write the concept 'house' than the Ge'ez script which needs only two characters. Teferi (2015:130) also presents the presence of some redundant alphabets of Geez alphabet (like ኣ, ዐ for /a/; ሀ, ሐ & ኀ for /h/; ጸ, ፀ for /ts/ and ሰ & ሱ for /s/) that represent the same sound with different character and he argues that this can be a good reason to reject the Ge'ez script. From the present research point of view, Teferi's (2015:130) argument is not convincing as the redundant alphabet stated above could be reduced to fit to the existing sounds of Oromo and the present Roman-base script of Oromo also undergone adaptation.

Another arena of the argument is the pedagogical issues if either Ge'ez or Roman based scripts are used to write the language. The first issue of this argument is about the number of the scripts which is related to pedagogy. Feyisa (1996:26) states that Oromo can be written using only 31 letters (*i.e., a, b, c, ch, d, dh, e, f, g, h, i, j, k, l, m, n, ny, o, p, ph, q, r, s, sh t, u, v, w, x, y, z,*) compared to 182 letters of Geez script. The implication of stating a number, here, by Feyisa

(1996:26) is that the Ge'ez script will take more time, money and energy to make someone literate than Roman-based script. But, this argument is counter argued by other scholars. Baye (1992:2) sees the case of number issues in two perspectives; from sound and script (kələm ቀለም). Baye (1992) keeps on explaining that there are not many differences between Roman based script and Ge'ez script if the basis of comparison is sound, then there are not many differences between Roman based script and Ge'ez script as the former has 26 and the later 27. If the perspective of comparison is a script (kələm ቀለም) a combination of vowels or vowels and consonants), again Baye (1992:2 claims that there is not much difference as Roman-based script has 130 (26 x 5) while Geez script has 189 (27X7). Baye (1992:3 also claims that in order to be literate in Rome-based script, some one has to learn at least two forms of each character (capital, small and the cursive form as in, for instance, A, a & *a* respectively).

The present researcher, however, argues that there is a significant difference when the parameters used for comparison is script (kələm ቀለም) as there are more number of scripts (kələm ቀለም) even when the count is taken as per Baye's count. The Ge'ez script has 59 more script (kələm ቀለም) than the Roman-based script. Regarding the forms of characters of Roman-based script, the present research claims that the forms (the capital, small and cursive) the characters will not need extra effort to learn as the three forms of some of the characters do not have much difference. Apart from this, the glottal sound /ʔ/ which is represented by the diacritic symbol (') and the frequently happening in the various texts of Oromo, is not considered by any of the scholars yet. Despite the glottal stop /ʔ/ is included in the phonemic inventory of Oromo by the standardization committee of Oromo; it is not officially represented with grapheme and included in the alphabet of Oromo.

The impact of choice of scripts to write Oromo in the process of literacy and the impact of multi-script on the country is also another issue that is to be considered in the argument. In this regard, Baye (1992:9) argues that having as minimum script as possible will benefit individual learner and a country as whole. From learner point of view, anyone who is literate in Roman-based script needs to be literate in Ge'ez script if one is to go and work in the other part of the country where the script is used (Baye 1992:9). The scholar further argues that such task of being literate

several times will definitely cost time, money, difficulty and inconvenience. The use of many scripts will transform Ethiopia from a multi-lingual to a multi-script one (Baye 1992:9). The scholar finally concludes that for multi-lingual under-developed countries like Ethiopia, it is uni-script and not multi-script that holds a clear advantage. To the contrary, Tilahun (1993:37); Wondimu (2019:24), Feyisa (1996:26) argue that the number of characters of Roman-based script is smaller than Ge'ez and therefore it is easier to learn. These authors further argue that the use of the Roman-based script would not put any additional burden on the children who have to learn English or want to learn other foreign languages which use the Latin alphabets. These authors who argue in favor of Roman-based script to write Oromo state that being literate in Roman-based script further helps learners to transfer skills to other languages of the world such as English, Germany, Swedish, Danish, French and others as these languages use the Latin script to write. Contrary to Baye's (1992) argument, Teferi (2015:203) presented counter argument and states that the use of a single script does not promote unity of the Ethiopian people. This is because that the use of only one script to write different languages in a country does not necessarily make the speakers feel as though they belong to a unified country. In this regard the present research also wants to emphasize that the use of a single script alone in a country may not unify the people of that country and the use of multi-script in a country too may not result in disintegration of a country from linguistic point of view.

There are also other parameters which are used to compare the Roman-based script and Ge'ez. The parameters are practical and attitudinal issues which are relevant for the choice of script to write Oromo. Regarding the parameter of practicality, (Feyisa 1996:26; Wondimu 2019:25) argue the Roman-based script has more advantage to write Oromo than the Ge'ez script as it is more easily adaptable to the available technology and could be easily used for commercial and telegraphic communication purposes. It is also argued from the proponents of Roman-based script that Ge'ez script is not able to represent consonant gemination and vowel as well as the implosive and the glottal stop. If Ge'ez script is used as an official script to write Oromo and special letters and diacritics are applied to represent the gemination, vowel length, implosive and the glottal stop, then, they may raise the cost and unavailability of printing equipment. Finally (Tilahun 1993:38, Feyisa 1996:26 Wondimu 2019:25) concluded that the adaptability of Roman-

based script to a computer technology and ease of learning is much more advantageous than the Ge'e Script.

According to Baye (1992:5), however, the entire problems of the practical issues rose by the proponents of Roman-based script for writing Oromo (such as the use of dot and diacritics) has nothing to do with linguistics; but it is basically a technological problem. The technological issues have to be resolved by the technologist and Ge'ez script can still be improved (adapted) to write Oromo. Moreover, Baye (1992:5) further argues that the case of consonant gemination and vowel length can be read (understood) from the context the word appeared without marking them or through other means mentioned earlier.

There are also non-linguistics issues of the arguments such as political and attitudinal. In this regard, Gadaa (1988:4) noted that the Roman-based script is relatively best suited to write Oromo as there are several political, cultural groups that strongly advocate for the use of the script. Mulualem (2017:6) also has strengthened the Gadaa's argument and states that Roman-based script is the symbolic representation and freedom or independency from past subjugation and domination of a political system. Wondimu (2019:18) & Mulualem (2017:6), finally, concluded that Roman-based script is a primary symbol (manifestation) and tool of unifying the Oromo people and should continue to be the writing script of the language. Baye (1992) claims that there is no linguistic reason to resort Roman-based script and political reasons should not be presented engulfed with linguistic one. He further states that if the Roman-based script, a script which is adapted from an exogenous Latin script and claimed by its proponent to be the symbol of its speaker, then an indigenous Ge'ez script, which is solely owned by Ethiopia, fully deserves to be the symbol of the country let alone the identity of a single nationality.

Finally, the present researcher argues that the present script (the Roman-based script or Qubee) should continue to be used for writing Oromo rather than any other script due to the following reasons. In a broader sense, there is no script or writing system that represents all aspects of a language and the Latin and Ge'ez scripts are not an exceptional (Baye 1992:4 and Kessler & Treiman 2015:15). Writing systems, therefore, only represent distinctions that result in lexical contrast. Regarding the Latin script, it is stated that great many modern languages use it, which,

in its original and most widely disseminated form, has no basic way of representing more than five different vowel qualities (Kessler & Treiman 2015:15). As a result, vowel length, stress, tone, or virtually any sound that was not found in Latin cannot be directly represented with the script. The authors further state that if the Latin script had had ways of representing these things, possibly many more languages today would as well. In this regard, neither Latin nor Ge'ez script is capable of directly representing vowel length and gemination which are major linguistic aspects of an argument on the choice of script.

Event though any script can be used with modification to write any language, the present research argues that the roman based script of Oromo should be continued to write Oromo. The reasons for such argument are due to the fact that the script has been used since 1991 (more than 28 years), the script is ratified to write the language and the attitude of the elites and users towards the script is strong (Muluaem 2017:9; Guluma 2018:7;Wondimu 2019:31). A great many people become literate using the Roman-based script or called "Qubee" and quitting using this script and replacing it with other script is almost impossible since replacing will make the literate one illiterate. Moreover, several books, newspaper, magazines have been published using Roman-based script (Qubee) which will be no more useful if the script is to be shifted to another script. This is in turn a destruction of resource of the country.

Roman based script (Qubee) helps for the better transparent representation of vowel length and consonant gemination by doubling the graphemes rather using either dot or 'sads' (6th consonant of the geminated sound and vowel) if Ge'ez is to be used to write Oromo. More than any other argument, ministry of culture and tourism has formulated a language policy of Ethiopian in 2020 which can be a solution to the argument. The policy states that the working languages of the federal government shall be Amharic, Oromo, Tigrigna, Somaligna and Afarigna (MoCT 2020: 14). This in itself indicates that the script of a working language of the federal government shall continue to be used to write the respective language. For instance, the fifth article of the constitution of the regional state of Oromia clearly states that Oromo shall be the working language of the regional state and it shall be written in the Latin alphabet (Constitution of Oromia regional state 1995:3). Moreover, the language policy of Ethiopia clearly stated in its preamble that accepting all languages and writing systems of the country as national heritages

and assets is fundamental principles of the policy (MoCT 2020:7).Furthermore, Ethiopian language policy, clearly stated in its section of protection of language rights that a nation, nationality or people have the right to choose writing system suitable for its language or to create a new writing system in compliance with appropriate standards (MoCT 2020:11).

The present research, therefore, argues that from the stated legal points of view, the already formulated official working Roman-based script (Qubee) of Oromo should continue to be used to write the language.

3.3.2.2 Development of Roman-based script (Qubee)

With the exception of using Oromo for the literacy campaign during the Derg regime, the pre EPRDF regimes did not formally declare a script to write Oromo and they did not use the language as a medium of instruction at any level. It was during the current government that the native languages were proclaimed to be used as medium of instruction and were studied as subjects of study. Thus, there was a need to develop one formal script that can be used by all speakers of the language uniformly.

It was in 1991 that the Roman based script or 'Qubee' became an official Oromo orthography which was adopted from the Roman based alphabet using the principle of writing through a one-to-one correspondence between sound segments and graphemes or symbols of the alphabet. The number of Oromo phonemes identified during the first official adaption was twenty- six. The phonemes [v]and [ts'] were not included in the phonemic inventory Oromo during the first adaption; however, currently, these phonemes are included by the standardization committee of Oromo.

Both Ge'ez and Roman based scripts were suggested during the discussion of which script to be used for official script of Oromo. But some of the reasons for the adoption of Roman based script for writing Oromo are the following. These are linguistic, pedagogical and practical reasons. As to Tilahun (1993:37), the linguistic reason stated during a meeting was that the Ge'ez script not only fails to indicate vowel length and gemination, but also slows down a writer's speed since

each symbol, which cannot be written cursively, must be printed. From the practical point of view the Roman based script was adapted to many languages of the world. Hence, an Oromo child who has learned Roman-based alphabet can learn the form of the English script in a relatively short period of time and its adaptability to computer technology can be taken as the merits of adopting the Roman based script.

The members of standardization committee have suggested about script choice during the focus group discussion that the current working Roman-based script of Oromo is appropriate to write Oromo as it is suitable to easily mark gemination and vowel length. The research participants further points out that it also enabled the language to grow faster as the script can be learned easily (see appendix G, FGD 1 24). The member of the committee also mentioned that the speakers did not claim that this script is not appropriate to write Oromo and this mean that they have accepted it (see appendix G, FGD 1 24).

Hence, there is an acceptance of the Roman based script to write the language from the standardization committee. This can strengthen the Tilahun's (1993:37) claim of writing the language using Roman based script. The constitution of the regional state of Oromia also states under its article five that Oromo shall be written in Roman based script (*Qubee*) (Revised constitution of Oromia regional state 2001:6). All in all, the current working script of Oromo, (Roman based script or *Qubee*) is now ratified by the Ethiopian language policy and is accepted by elite of Oromo (3.3.2.1).

3.4. Punctuation in Oromo

Punctuation marks used in both Oromo and English are the same and used for the same purpose with the exception of the apostrophe. Apostrophe mark (‘) in English shows possession, but in Oromo it is used in writing to represent the glottal stop (called '*hudhaa*').

Table 8. Oromo Punctuations

Symbol	name of punctuation	mark	gloss
.	<i>tuqaa</i>		full stop
?	<i>mallattoo gaaffii</i>		question mark
,	<i>qodduu</i>		comma
;	<i>qodduu jabaa</i>		semi colon
:	<i>tuqlamee</i>		colon
-	<i>taraarii</i>		hyphen
!	<i>diniqa ykn raajeffannoo</i>		exclamation
()	<i>gaaffee</i>		parenthesis
" "	<i>waraabee lakkuu</i>		quotation mark
...	<i>fufii</i>		ellipsis
' '	<i>waraabee qeenxee</i>		single quotation

Oromo punctuation marks follow the same punctuation pattern used in English. The full stop (.) in statement, the question mark (?) in interrogative and the exclamation mark (!) in command and exclamatory sentences marks at the end of a sentence, and comma (,) which separates listing in a sentence and the semi colon (;) listing of ideas, concepts, names, items, etc. The other quotation marks are also used just like in English. The double quotations (“ ”) and single quotation (‘ ’) are used in similar rules as in English. Regarding its application in the corpus this study consulted, except for *qodduu jabaa* ‘semi colon (;), *fufii* ‘ellipsis’ (...) and *qodduu* ‘comma’ (,), all the rest punctuation marks are consistently applied in corpuss. Semi colon (;) and comma (,) are seen to be replacing one another in few case in the corpus this study consulted. On the other hand, *fufii* ‘ellipsis’ (...) is also seen to be replaced by four dots (....) in the primary and secondary school textbooks (MoE grade 12 2012:133) rather than three dots.

3.5 Numerals

In Oromo orthography, Arabic numerals are used to write numbers and the numerals can be represented and written in words, cardinals, in ordinals and fractions. To show how numbers are represented in Oromo orthography, some illustrations are given in the following table.

Table 9. Oromo cardinals and their Arabic equivalent

Cardinal numerals in words	Arabic	cardinal numerals in words	Arabic	cardinal numerals in words	Arabic
<i>tokko</i>	1	<i>kudha tokko</i>	11	<i>soddoma</i>	30
<i>lama</i>	2	<i>kudha lama</i>	12	<i>afurtama</i>	40
<i>sadii</i>	3	<i>kudha sadi</i>	13	<i>shantama</i>	50
<i>afur</i>	4	<i>kudha afur</i>	14	<i>jaatama</i>	60
<i>shan</i>	5	<i>kudha shan</i>	15	<i>torbaatama</i>	70
<i>ja'a</i>	6	<i>kudha ja'a</i>	16	<i>saddetama</i>	80
<i>torba</i>	7	<i>kudha torba</i>	17	<i>sagaltama</i>	90
<i>saddet</i>	8	<i>kudha saddeet</i>	18	<i>dhibba tokko</i>	100
<i>sagal</i>	9	<i>kudha sagal</i>	19	<i>kuma tokko</i>	1,000
<i>kudhan</i>	10	<i>digdama</i>	20	<i>miilliyona tokko</i>	1,000,000

There are various ways of representing ordinals in Oromo. The most frequently used Oromo ordinal and other possible ordinals in Oromo are shown in the following consecutive tables.

Table 10. The most frequently used Oromo ordinals and their short forms

Ordinals in words	ordinals in short form	gloss
<i>tokkoffaa</i>	<i>1^{ffaa}</i>	1 st
<i>lammaffaa</i>	<i>2^{ffaa}</i>	2 nd
<i>saddaffaa</i>	<i>3^{ffaa}</i>	3 rd
<i>afraffaa</i>	<i>4^{ffaa}</i>	4 th
<i>shanaffaa</i>	<i>5^{ffaa}</i>	5 th
<i>ja'affaa</i>	<i>6^{ffaa}</i>	6 th
<i>torbaffaa</i>	<i>7^{ffaa}</i>	7 th
<i>saddettaffaa</i>	<i>8^{ffaa}</i>	8 th
<i>saglaaffaa</i>	<i>9^{ffaa}</i>	9 th
<i>kurnaffaa</i>	<i>10^{ffaa}</i>	10 th

Although ordinal numeral stated in the above table is the one that is most frequently used in the various Oromo text and reference books, there are various ways of representing the ordinal form of Oromo. For instance *-eeysa/-eeysoo/-eessaa /-eessoo/-eechaa/-eechoo/ - affaa/-affoo/* are also the possible representation of ordinal numerals in Oromo orthography (Abdulsamad 1994:81). Their respective possible short forms for the first order (1st), for instance, are *I^{ysa}, I^{ysoo}, I^{ssaa} I^{ssoo}, I^{chaa}, I^{choo}, I^{ffaa}, I^{ffoo}*. The following table shows the possibilities of writing ordinals in Oromo. All ordinals in the language can be represented in the following ways; but only the first and second ordinal numbers in Oromo are given as an illustration.

Table 11. Variations in Oromo ordinals and their short forms

<i>-eeysa</i>	<i>-eeysoo</i>	<i>-eessaa</i>	<i>-eessoo</i>	<i>-eechaa</i>	<i>-eechoo/</i>	<i>-affaa/</i>	<i>-ffoo</i>	glo ss
<i>tokkee- ysa</i>	<i>tokkeey- so</i>	<i>tokkeessaa</i>	<i>tokkeesso</i>	<i>tokkeech- aa</i>	<i>tokkeech -oo</i>	<i>tokkoffaa</i>	<i>tokko- ffoo</i>	
<i>I^{ysa}</i>	<i>I^{ysoo}</i>	<i>I^{ssaa}</i>	<i>I^{ssoo}</i>	<i>I^{chaa}</i>	<i>I^{choo}</i>	<i>I^{ffaa}</i>	<i>I^{ffoo}</i>	1 st
<i>lammee- ysa</i>	<i>lammee- yso</i>	<i>lammeessaa</i>	<i>lammeess -oo</i>	<i>lammeech- ha</i>	<i>lammee- choo</i>	<i>lammaffaa</i>	<i>lammaffo -o</i>	
<i>2^{ysa}</i>	<i>2^{yso}</i>	<i>2^{ssaa}</i>	<i>2^{ssoo}</i>	<i>2^{chaa}</i>	<i>2^{choo}</i>	<i>2^{ffaa}</i>	<i>2^{ffoo}</i>	2 nd

As we can see in table 11, Oromo can have eight ordinal and their respective short forms. Still the ordinals stated in the above table can be made to be more shorter; like *I^{ysa} → I^{sa} Ist, I^{ysoo} → I^{so}, I^{ssaa} → I^{ssa}, I^{ssoo} → I^{ssoo}, I^{chaa} → I^{cha}, I^{choo} → I^{cho}, I^{ffoo} → I^{fo} and I^{ffaa} → I^{fa}*. Although the ordinal '*ffaa*' is frequently used in textbooks and in various literature; including in those written material compiled by the Oromo standardization committee and published by Oromia culture and tourism bureau, the committee did not clearly selected and normalized any of the above short forms of Oromo ordinals. On the other hand, as it can be observed in various reading materials published by Oromia culture and tourism bureau, Oromo ordinal is not used similarly. For instance, in a grammar book published by the bureau, the short form ordinal '*fa*' is used (Qajeelcha 1998:254, 262, 263). This shows, this concerned body of standardizing Oromo also

applied another form of the short form of the ordinal '-ffaa' (Qajeelcha 2008:151). Hence, selecting and harmonizing one of the above ordinals is very crucial. The ordinal '-ffaa' is currently the most frequently used ordinal in Oromo and still it can be modified and regularized as follows.

Table 12: Recommended ordinal short forms of Oromo

Oromo ordinal in words	the current short form	recommended short form	gloss
<i>tokkoffaa</i>	<i>1^{ffaa}</i>	<i>1^{fa}</i>	1 st
<i>lammaffaa</i>	<i>2^{ffaa}</i>	<i>2^{fa}</i>	2 nd
<i>saddaffaa</i>	<i>3^{ffaa}</i>	<i>3^{fa}</i>	3 rd
<i>afraffaa</i>	<i>4^{ffaa}</i>	<i>4^{fa}</i>	4 th
<i>shanaffaa</i>	<i>5^{ffaa}</i>	<i>5^{fa}</i>	5 th
<i>ja'affaa</i>	<i>6^{ffaa}</i>	<i>6^{fa}</i>	6 th
<i>torbaffaa</i>	<i>7^{ffaa}</i>	<i>7^{fa}</i>	7 th
<i>saddettaffaa</i>	<i>8^{ffaa}</i>	<i>8^{fa}</i>	8 th
<i>saglaaffaa</i>	<i>9^{ffaa}</i>	<i>9^{fa}</i>	9 th
<i>kurnaffaa</i>	<i>10^{ffaa}</i>	<i>10^{fa}</i>	10 th

As we can see in table 12, since there is gemination and vowel length in the short form of the ordinal *-ffaa*, it can still be modified and normalized as *1^{ffaa} '1st' → 1^{fa} '1st'*. But, since the main objective of opting the Roman-based script to write Oromo is to transparently mark consonant germination and vowel length, this research argues that *-ffaa* should be codified to represent Oromo ordinal.

3.6 Spelling convention

All of the following data under each convention are collected from Qajeelch (1998), grade 7 Biology textbook (MoE 2013) and from Qajeelcha Wiirtuu volume 1-5 2008:26).

1. Words in a sentence are separated from each other using space. For instance:

(2) *Inni kitaaba siif kenne*
Inni kitaaba sii-(f) kenn-Ø-e
 He:Nom book you-DAT give-3sm-PV
 'He gave you book'

2. The first letter of every word in the beginning of a sentence should be written in capital letter. The initial letter of names of countries, proper nouns, names of oceans, places and rivers should be written in capital letters in any position of a sentence.

3. Vowel length and consonant gemination are phonemic in Oromo and they are marked in the orthography of the language. Hence, short phonemes are written with a single grapheme whereas long phonemes are written by doubling the graphemes to show vowel length and consonant gemination. Unlike the case of Hadiyya, in which geminations of digraphs are marked by repeating the entire digraph, gemination of digraphs in Oromo is not marked.

4. The copula 'dha' shall be suffixed and shall not be written alone.

(3) a) *Caalaan barataa cimaadha.*
Caalaa-n barataa cimaa-dha.
 Chala-NOM student clever- COP
 'Chala is a clever student'

b) **Caalaan barataa cimaadha.*
Caalaa-n barataa cimaa dha
 Chala-NOM student clever COP
 'Chala is a clever student'

5. The conjunction *fi* 'and' shall be suffixed and shall not be written alone

As an illustrative examples are given as follows.

(4) Accepted	not accepted	gloss
<i>kitaabafi qalama</i>	* <i>kitaaba fi qalama</i>	'book and pen'
book.and pen	book and pen	

Although, the standardization committee of Oromo set rules regarding how the Oromo conjunction shall be written, it is being spelled sometimes attached and in some other cases detached from a word before it (Qajeelcha 1998:82). Hence, this is a failor of the application of conventions set by the committee.

6. Most words formed by reduplication in Oromo are guided by rule as we can see in the following examples.

(5)	Reduplication	gloss
a)	<i>cabaa + cabaa → caccabaa</i>	'group of broken'
	<i>cimaa + cimaa → ciccimmaa</i>	'group of clever'
	<i>dheraa + dheraa → dhedheeraa</i>	'group of long'
	<i>dimaa + dimaa → diddimaa</i>	'group of red'
	<i>gabaabaa + gabaabaa → gaggabaabaa</i>	'group of short'
	<i>jabaa + jabaa → jajjabaa</i>	'group of strong', 'youngers'
b)	<i>adii + adii → adaadii</i>	'group of white'

In Oromo, words that are formed by reduplication, following the beginning consonant and vowel of the first word, the first consonant of the first word will be geminated and first vowel of the first word will follow the geminated consonant and the new word formed by reduplication will be written as one word without any hyphen. Most reduplicated words are formed and written like the one we have seen in the above illustration. In other words, the whole root will be reduplicated. But there are also exceptions as in data 5b.

7. Oromo has its own rule of writing loan words. Accordingly loanwords are written via adapting into the Oromo phonological system. The following data taken from Qajeelcha 2008) and grade 7 Biology textbook (MoE 2013) shows how loan words are adapted into the phonological system of the language.

(6) Oromo	gloss	Oromo	gloss
<i>baankii</i>	'bank'	<i>koolleejii</i>	'college'
<i>ispoortii</i>	'sport'	<i>iskoolaarshiippii</i>	'scholarship'
<i>kompitara</i>	'computer'	<i>tiraaktara</i>	'tractor'
<i>akroobaatii</i>	'acrobat'	<i>kitfoo</i>	'kitfo' or 'minced meat'
<i>kaameraa</i>	'camera'	<i>kaabboo</i>	'capo' or 'foreman'

Various phonological adjustments are made when Oromo borrowed words from various languages. As the data 6 shows, when the borrowed words end in a vowel in their original language, the final vowel will become longer as in the of '*kitfo*'→'*kitfoo*' and '*capo*' →'*kaabboo*'. On the other hand, when the language borrowed words which begin with consonant cluster, the words will be integrated into Oromo by inserting the vowel (<*i*>) before the initial consonant cluster or in between consonant cluster in order to adhere to the rule that no Oromo words begin with consonant cluster as in, for instance, *ispoortii* 'sport' and *tiraaktara* 'tractor' respectively. Word final consonant clusters are also adapted via the addition of either the vowel <*i*> or <*a*> or their respective long vowel. But the committee still did not codify which vowel to add when.

8.As per the decision of the standardization committee of Oromo, when a word begins with vowel length, the glottal stop must be included in the initial part of the word as in '*aadaa*' 'culture', '*aanaa*' 'assistant' etc. (Qajeelcha 2008:26). In fact, no Oromo words begin with a vowel. But, this is not seen implemented in the various publications, including in the publications of the committee itself. Hence, the committee finally, reversed the later convention to the previous convention (the application of a glottal stop <'> only word medially as in, for instance, *bu'aa* 'result'.

There are still other word break issues which the standardization committee of Oromo did not give due attention. For instance, there is no convention as to how elements like *wal*, *haa*, *ni*, *kan*, *nan* should be written. The elements are observed to be written sometimes attached to words that appear before or after them. In other instances they are being written as independent words.

Though word break rules are generally conventional, it has to be supported by scientific linguistic criteria. In this regard, Lojenga (2014b: 91-99) provides a good set of criteria for word boundary placement. The author identifies three types of morphemes: independent word, a clitic or a bound morpheme. The three criteria for the identification of words are syntactic, phonological, and semantic. The syntactic criteria itself can be seen from the perspective of mobility (the ability for a morpheme to be in different places in a sentence), separability (the morpheme can separate from the neighbouring lexical morpheme by the insertion of another morpheme) and from substitutability (a lexeme can be substituted by a grammatical morpheme in the same syntactic slot). The phonological criteria can also be subdivided into pronounceability in isolation, phonological unity (influence from the lexical morpheme to grammatical morphemes and/or vice versa by phonological processes) and phonological bridging (how to divide words in a phrase or sentence).

Coming back to our concern, the current word break issues being practiced in Oromo written language is a bit problematic as it has been mentioned above. It needs due attention and a detailed study has to be conducted. Tamam (2015:83) states that disambiguation rather than ‘we write as we speak’ motto should be followed. He further points out that generally the preferences of writing certain morphemes, mainly clitics, disjunctively have been stressed with few suggested to be conjoined with hyphen. The present study proposes, based on the criteria mentioned above, these elements to be written attached to the words that appear before or after them as the elements do not have meanings independently. Hence, like any other functional elements in the language they better written together with the preceding or after them.

3.7 Evaluation of Oromo orthography

This section discusses the evaluation of Oromo orthography in the light of standardization and orthography development principles in the following consecutive subsections. Smalley (1963:31–52) states a list of five factors that summarize the important aspects of developing and accessing a good orthography. Smalley’s five criteria for developing a good orthography are presented in order of importance as follows.

1. Maximum motivation for the learner and acceptance by the speech community
2. Maximum representation of speech

3. Maximum ease of learning
4. Maximum transfer of skills
5. Maximum ease of reproduction

Thus, Oromo orthography will be evaluated in light of orthographic and standardization principles stated above and with other related concepts discussed under the second chapter.

3.7.1 Maximum motivation for the learner and acceptance by the speech community

Smalley (1963:31–52) identifies three categories of people who have a stake in the design of orthography, and who need to be motivated towards the support of orthography: the national government, the individual learner, and the speech community. Some governments take an active role to influence orthography development. Their interest is that the orthographies serve the ultimate purpose of uniting their people and creating national identity.

Individuals and speech communities are motivated by both their reasons for wanting orthography and their perceptions as to how their orthography should look. The motivations of an individual greatly overlap with the general motivations of the wider speech community. The individual needs a purpose for expending the effort to learn to read and write. The purpose may be for education and access to more knowledge and information, or for greater access to economic benefits. Some people have an interest in maintaining a sentimental connection with their heritage language for maintaining an identity, particularly if it is endangered. Some people may want to relearn the language of their ancestors, or they may be motivated as an expression of religious devotion (Decker 2014:31).

As it has been discussed in the previous sections, Oromo was written using various scripts by various authors and bodies beginning from early 1840s up to 1990 (see section 3.3.1). The language has also been written using Ge'ez script during the Derg regime for the purpose of literacy campaign as using other scripts was not officially encouraged. But from 1991 onwards it has been written using Roman-based script. Regarding the appropriateness of Roman-based script to write Oromo, points have been mentioned during the FGDs as follows

What we know is that the current Roman based script of Oromo is appropriate for Oromo. It also enabled the language to grow faster. Speaker of the language did not ask that this script is not appropriate to write Oromo. Hence, this can mean that they have accepted it. If there is script and orthographic related problems a research has to be conducted and presented to the government and to the speakers (speech community) as a whole to be accepted

(See Appendix G, FGD 1, 24).

In addition to the ratification of the Roman-based script to write Oromo by Oromia regional government constitution under article five (Revised Constitution of the Oromia National Regional State 2001:6), it is also mentioned during the focus group discussion that the Roman based script is more appropriate to write Oromo. The above data also asserts that there were no objections from the speech community side regarding the current working script. This indirectly shows that the speech communities have accepted it. Apart from this, it is for its suitability to easily and transparently mark the basic features of Oromo that the script is chosen. Accordingly, in addition to symbol sound correspondence, phonological features such as vowel length and gemination are marked for the reason that priority is given for transparency than economy and this case is also true for Hadiyya (Samuel 2017: 110).

Oromo is being taught in primary school level (grade 1- 8) as medium of instruction and as a subject from grade 1-12. It is also being given as a subject from primary to university. The language is also currently being used in the various mass media as a medium of transmission, even in the foreign media like on the British Broadcasting Corporation (BBC World). Hence, we can conclude that Oromo is now being used in all domains; particularly in the educational, administrative and media domains.

3.7.2. Maximum representation of speech

Maximum representation of speech refers to the relationship between the graphemes and the sounds of a language. It is the maximum possible representation of sounds. The concerned body which is given a power legally to standardize Oromo claims the number of Oromo phonemes to be twenty eight (Qajeelcha 1998:14). But only twenty six of these are represented by grapheme

and included in the Oromo alphabet. The alveolar ejective, /s'/ is included in the phonemic inventory of the language. But, though this phoneme is represented with digraph 'ts', it is not included in the Oromo alphabet. The representation of the phoneme /s'/ with grapheme <ts> in the phonemic inventory of the language have reasoned out that the phoneme has been introduced to Oromo via Amharic loanwords, as in <tsabala>. This concerned body to standardize the language further claims that they have included the phoneme /s'/ into Oromo phonemic inventory stating that this phoneme entered into Oromo through loanwords. But, Oromo loanwords with the phoneme /s'/ and which is being written with grapheme <ts> are rare. Moreover, loanwords that comprise [s'] could be written through adaptation to the nearest existing phonemes /t'/ as <xabala> 'Holly water'. The above legal body did not consider and made any decision regarding the phoneme /ʒ/ which entered in Oromo through loan words like television. The word 'television' is being used and written many times daily; particularly in the media sphere. But this word is being written differently as <televiizhiinii>, <televiizyiinii> and <televiijiniini> due to different representation of the phoneme /ʒ/ as /Zh/, /Zy/ and /J/. Concerning these phoneme, it is mentioned during the FGDs that the loan phoneme entered through loanwords and various individuals and mass media have represented this phoneme differently as we did not still either represented the phoneme with grapheme or made any decision officially (see appendix G, FGD 1, 29). The research participant finally concluded that this has to be studied and either has to be represented with a grapheme or has to be clearly stated as to how the phoneme /ʒ/ should be written in the future.

On the other hand, the unsettled issues about the graphematic representation of the glottal stop or sound represented by /ʔ/ in the IPA inventory should be one of orthographic issues to be discussed. This study observed that the phoneme occurrences frequently in the texts this study consulted, particularly word internally. Although this phoneme is included in the phonemic inventory of the language, the grapheme for this phoneme is not still devised and included in the orthography of the language. Contrary to the phonemic inventory of the consonants of Oromo and their representation by graphemes, the glottal stop or 'hudhaa' with the phonemic status in the language is indicated in the orthography of Oromo by the English apostrophe <'>. It is not included in the list of graphemes of the language. It is mentioned during FGD about the phoneme as follows.

On the other hand, though glottal stop /ʔ/ has a phonemic status and included in the phonemic inventory of the language, the phoneme is not yet represented with a grapheme and it is not included in the alphabet of the language. This has to be studied and we are the concerned body to study it. The glottal stop is one of the frequently occurring phonemes in the orthography of the language which is being indicated by apostrophe ('). Since we have included in the phoneme in the phonemic inventory of the language, we have to study and find out how it could be represented using by grapheme. Some times this phoneme is also being seen replaced and written by 'w, y and h'. Standardizing the language takes several years. So, this has to be studied and the solution has to be given legally.

(See Appendix G, FGD 1, 30).

The above data asserts what has been mentioned before the quote. Despite its frequent occurrence (at least in the text corpus that this study consulted) and its inclusion in the phonemic inventory of the language, the concerned body did not give due attention to devise an appropriate grapheme for the phoneme. The above research participant stated that it has to be studied and the standardization committee of the language is the concerned body to study and decide on the issue. The following data collected from Wiirtuu volume 10 illustrative examples show the phonemic status of the glottal stop /ʔ/.

(7) Orthography	IPA	gloss
a) <i>ba'e</i>	<i>baʔe</i>	'he went out'
<i>bare</i>	<i>bare</i>	'he aware (know)'
b) <i>sa'a</i>	<i>saʔa</i>	'cow'
<i>sana</i>	<i>sana</i>	'that'
c) <i>dha'e</i>	<i>dʰaʔe</i>	'he struck'
<i>dhare</i>	<i>dʰare</i>	'he lied'
d) <i>fe'e</i>	<i>feʔe</i>	'he loaded'
<i>fedhe</i>	<i>feʔe</i>	'he liked'
e) <i>taa'e</i>	<i>taaʔe</i>	'he sit down'
<i>taahe</i>	<i>taahe</i>	'he sit down'
<i>taaye</i>	<i>taaye</i>	'he sit down'
f) <i>tamsa'uu</i>	<i>tamsaʔuu</i>	'transmit'
<i>tamsawuu</i>	<i>tamsawuu</i>	'transmit'
g) <i>arma</i>	<i>arma</i>	'here'
<i>harma</i>	<i>harma</i>	'breast'

As it depicted in the above data, the minimal pairs presented in (7a) up to (7d) shows the phonemic status of the glottal stop /ʔ/ in the language. On the other hand, the list of data presented in 7e and 7f shows that the glottal stop can sometimes be replaced by /w/, /h/ and /j/ as it has been also stated in the previous quote before illustration seven. The data under 7g show the word initial occurrence of glottal stop. The standardization committee of Oromo decided two times regarding the representation of the glottal stop /ʔ/. The first decision was to consistently omit in word initial position and marking only word medially. The second decision was the representation of the glottal stop /ʔ/ word initially when a word began with long vowel as in *ʔaadaa* <'aadaa> 'culture' which was previously known to be written as <aadaa> as well as word medially as in *sa'a* 'cow' (Qajeelcha 2008:1). Finally, the committee reversed the second decision to the first decision, i.e., marking the glottal stop only word medially as it is stated in Qajeelcha (2008:14) just like the case of Hadiyya (Samuel 2017:111). The reversal of the

decision of marking the glottal stop word initially could be due to the confusion that it creates with single punctuation mark as seen in (8) below.

- (8) *"aadaa kabajuun gaaridha' jedhe Tullun.*
"aadaa kabaj-uu-n gaarii-dha' jedh-e Tullu-n
Culturerespect-VNM-NOM good-COP said-PV Tullu-NOM
'respecting culture is good' said Tullu.

Here, from (8), we note that there occurred the confusion between the single quote and the glottal stop marker (apostrophe) or simply (') that may challenge at least learning to identify it from a single quote (') at an early stage. Hence, the first decision of the committee, i.e., consistently using it at the word medial and through avoiding when it occurs at the beginning of a word is a wise decision so as to avoid the issue of confusion with punctuation marks. But, still glottal stop is phonemic word internally in Oromo, this study recommends the IPA symbol ʔ for grapheme representation of the phoneme and it could be placed on the final part in the alphabetic order of the language so as to avoid confusion with the first letter of the language <a>. The grapheme <ʔ> can easily be written cursively and be represented in a computer via the use of software. This recommendation is due to three reasons. First, the standardization committee of Oromo codified the glottal stop as a phonemic sound and included in the phonemic inventory of the language (Qajeelcha 1998:14). Second, if it is phonemic sound, then a grapheme has to be designed just like any other phonemic sounds of the language rather than representing it with a punctuation marker ('). Here, of course, Oromo is now considering the punctuation mark (') as grapheme to represent the glottal stop. But, this sign is representing two concepts; the right single quotation mark as well as the glottal stop. Third, the grapheme of the sound has to appear in the alphabet order of the language. Otherwise, there is no reason of representing it with a punctuation mark (') and not including it in the alphabetic order of the orthography of the language. As a general principle, a sound which contrasts in a language should be represented by a different letter. That is, if separate phoneme is established, this needs separate grapheme, not punctuation or diacritic markers (Cahill & Karan 2008:4).

From the general principle of grapheme design point of view, diacritic and punctuation marks are not recommended to represent the basic sounds of a language as we can see from the process of establishment of the script and orthography of Somali language. Here, stating a few of the most important principles that the 1960 Somali Language Committee strictly followed successfully seems important for the issues of grapheme representation of the Oromo glottal stop. As stated by the committee, the script and orthography of Somali language must not have any diacritic or punctuation mark except by those which have been approved by the committee, i.e. one for accent and one to modify sound and must not have any sign which has more than one function (Banti and Ismail 2012:37). It is also stated and strictly implemented by the committee that the script and orthography of Somali language must not have signs which are differentiated only by means of diacritics or punctuation mark. The committee also mentions that the script and orthography of the language must not have any punctuation or diacritics which are representing basic sounds of the language themselves.

There is also an unsettled issue of the underrepresentation of digraphs in their geminated form of realization. To this end, all phonemes represented by digraphs are currently being written using the non geminated form in the orthography of Oromo whether they are realized as geminated or non-geminated. Since the issue of the representation of gemination has been one of the factors for the decision to adopt the Roman based script for writing Oromo, it is vital that careful attention is given to it with a view to making it simpler and more economical to use. In this regard, it is only <ch> /tʃ/ which is always realized in geminated form. Hence, there is no need to distinguish the non-geminated from the geminate form as a result of the phoneme's realization in only geminated forms regardless of the position in the word it occurs or how it is articulated. Despite the fact that the remaining four phonemes (/d/, /pʰ/, /j/ and /ʃ/) are realized as both geminated and non-geminated, they are not being marked to distinguish the geminated from the non-geminated. Of course, the frequency of occurrence of these phonemes as geminated forms is not as frequent as other phonemes of the language. The following are collected from Wiirtuu volume 4 1991 and Qajeelcha (1998) as an illustration to show the existence of the above phonemes in both geminated and non-geminated realizations.

(9)	Oromo	IPA	gloss
a)	<i>hadhaa</i>	<i>hadɗaa</i>	'poison'
	<i>hadhaa</i>	<i>hadaa</i>	'cleaner'
	<i>hodhuu</i>	<i>hodiuu</i>	'to suck'
	<i>hodhuu</i>	<i>hodɗuu</i>	'to sew'
	<i>xinnaadhaa</i>	<i>t'innaadɗaa</i>	'you become lower status'
b)	<i>addunyaa</i>	<i>ʔadduɲɲaa</i>	'world'
	<i>fakkeenya</i>	<i>fakkeɛɲɲa</i>	'example'
	<i>funyaan</i>	<i>fuyɲaan</i>	'nose'
	<i>funyoo</i>	<i>fuyɲoo</i>	'rope'
	<i>nyaata</i>	<i>ɲaata</i>	'food'
c)	<i>ashamaa</i>	<i>aʃʃamaa</i>	'greetings'
	<i>bashannana</i>	<i>bafaannana</i>	'entertainment'
	<i>shaashamanne</i>	<i>eʃaʃamanne</i>	'name of a town'
d)	<i>happeessa</i>	<i>hap'p'eessa</i>	'adhesion'
	<i>suphamaa</i>	<i>sup'p'amaa</i>	'sewed closes'

As the above data shows, the four phonemes (/d/, /p'/, /ɲ/ and /ʃ/) are realized both as geminated and non-geminated forms. But the occurrence of the phoneme /ʃ/ in its geminated form is rare. The geminated and the non-geminated occurrences of these phonemes, particularly the first three phonemes, are not currently being distinguished through marking the geminated one in the orthography of Oromo. It is true that the contexts in which the geminated and the non-geminated contrasts are not many as compared to other phonemes. Of course extra study (evidence) is also needed to check whether the lack of marking the geminated realization of the phonemes can create reading and understanding problems to users or not. But, the orthography will be more transparent if they are marked. To this end, the standardization committee didn't decide as to how these phonemes should be written to clearly differentiate their geminated form from the non-geminated one.

On the other hand, there is also a standardization gap regarding how the geminated forms of these phonemes are represented by digraphs are to be written. This study, therefore, recommends that the geminated realization of these sounds has to be marked. If they are to be marked, there

are two options to do. The first option is doubling only the first constituent of the digraphs as in, <hodddhuu> 'to sew'. The second option is doubling both constituents of the digraphs as in, <hodhdhuu> 'to sew' just the case of Hadiyya in which all the digraphs in the language (ch, sh, ph, zh, ts, and ny) is decided by the concerned body to be written this way to mark gemination (see Samuel 2017:103). This study believes that such means of marking geminated realization of digraph can even be worse than the previous problem of (unmarkedness) as doubling the entire digraph is not economical. Moreover, the second option gives more emphasis to transparency than economy. Hence, although transparency is the major issue for which Oromo was decided to be written using the Roman based script, marking the geminated realization of the digraphs using the second option stated above is not appropriate and not economical. So, in order to reconcile transparency (to be marked) and economy, this study recommends the doubling of only the first constituent of the digraphs as in <hodddhuu> 'to sew'. Here, doubling <d> is better than doubling <h> as /d/ and /d/ are both alveolars and /d/ will assimilate /d/ when it occurs before it.

3.7.3. Maximum ease of learning

Ease of learning can be motivational and is greatly impacted by the orthography's relationship to the spoken language. A good orthography that encodes the information needed by the mother tongue reader obviously facilitates ease of learning. The reduction of ambiguity and keeping the simplicity of a writing system is important to enhance its clarity (Stark 2010:12). To this end, all Oromo graphemes are necessary to write Oromo. But, though the phoneme is not frequently occurring, the issues of /ʒ/ in loanwords of Oromo has to be decided about the representation of the phoneme as currently the phoneme is being represented <Zh>, <Zy> and with nearest phoneme /dʒ/<J> particularly in the media sphere.

Task of standardization of orthography does not end in only normalizations of graphemes, marking geminated realizations of phonemes and the regularization alphabet order of orthography of a language. The standardization of orthography of a language requires the normalization of word breaking issues. Maintaining the clarity of orthography through utilizing consistent way of word breaking like how each compound, pronouns, morphemes and conjunctions can highly affect ease of learning. From this point of view, there are a number of

variations in the language. This issue has also been confirmed during FGD. The research participants state that there are wide orthographic variations in the textbooks of primary school and even in the materials the committee compiled as the texts are written without looking at what has been codified (See Appendix G, FGD 1, 57). The idea put forwarded during the focus group discussion asserts the existence of a wide range of orthographic variations, particularly with regard to word divisions and spacing of compounds, affixes and conjunctions. For instance, <fi> 'and', <dha> 'copula', <ti> 'copula or genitive' are decided to be written attached to the words before them as they are pronounced as part of the word before them (See Appendix G, FGD 1, 57). But they are being written sometimes attached and sometimes detached from word occurring before or after them. It has also been mentioned during the discussion that the committee is going to work hard to minimize such kind of variation as the right and responsibility is given to them legally to standardize the language (See Appendix G, FGD 1, 57). The following data in 10a, 10b, 10d & 10e are collected from grade 7 and 8 Oromo textbooks and the data in 10c & 10f are collected from Qajeelcha 2008 of the standardization committee of Oromo to show some of such variations.

- | | |
|--|--|
| <p>(10) a) <i>Anaafi isa</i>
 <i>Anaa-fi isa</i>
 I:and he
 'I and he'</p> | <p>d) <i>Anaa fi isa</i>
 <i>Anaa fi isa</i>
 I and he
 'I and he'</p> |
| <p>b) <i>Kun mana koo-ti</i>
 <i>Kun mana koo-ti</i>
 House my-COP
 'This is my house'</p> | <p>e) <i>Kun mana koo ti</i>
 <i>Kun mana koo ti</i>
 this housemy COP
 'this is my house'</p> |
| <p>c) <i>Kun garbuu-dha</i>
 <i>Kun garbuu-dha</i>
 this barley-COP
 'This is barely'</p> | <p>f) <i>Kun garbuu dha</i>
 <i>Kun garbuu dha</i>
 this barely COP
 'this is barely'</p> |

As data (10) show, orthographic variation of elements is being observed in the primary school textbooks and even in the various publications of the standardization committee of Oromo. In order to minimize such variation, the standardization committee of Oromo codified the elements

ti and *dha* stated above to be written attached to words before them. Hence, all illustrations under (10a), (10b) and (10c) are the codified and right norm and those under (10d, 10e and 10f) are 'unacceptable'. There is also orthographic variation as to how the prefixes of the language are written. For instance, the particle *haa* (jussive) and the element *hin* 'negation' are currently being written inconsistently. Sometimes they are written attached to the word after it as in <*ani hin-nyaann-e*> (I NEG-eat.IPV 'I did not eat' and in some other cases they are written as an independent word as in <*ani hin nyaanne*> 'I did not eat'. In most cases these elements are observed to be written as independent words in the text corpus this study consulted. But the standardization committee of Oromo codified it to be written attached to the word after it. One can note from this analysis that what is being codified is not put in practice. Such challenges may negatively affect the standardization process of the language.

3.7.4 Maximum transfer of skill

Maximum transfer refers to the relationship of the orthography of a non-dominant speech community to the orthography of the national or regional language. Gralow (1981:10) points out that people in non-dominant speech communities will eventually need to transfer their reading skills to the national language. Once people gain literacy in their heritage language, they should be able to shift to literacy in another language with as little difficulty as possible. It follows that primary literacy in a person's heritage language aids in the acquisition of literacy in another language (Decker 2014:56, Thomas & Collier 2002: 20). Regarding maximum transfer of skill, to the speakers of Oromo, it is important to learn Amharic and English so as to access wider opportunity in the country. In the case of Oromo, students are expected to learn Amharic script to be literate in Amharic as there is no relationship between the Roman and Ethiopic scripts.

The graphemes which exist in both Oromo and English, like <*c*>, <*q*>, <*x*>, and <*ph*>, but with different phonetic value (a single grapheme representing different sounds inter and intra languages) may have a negative impact on the transfer of skills during early grade reading of English for Oromo mother tongue learner though it requires further study. The sounds of these graphemes are ejectives in Oromo as in *c* [t], *q* [k'], *x* [t'] and *ph* [p']. These ejectives are existent in Oromo and non-existent in English phonology and such type of sound discrepancy may negatively affect the transfer of skills. Hence, the present study recommends replacement of the

existing graphemes of the ejectives c [tʃ], q [kʰ], x [tʰ] and ph [pʰ] with a new grapheme <c'>, <k'>, <t'> and <p'> and marking their geminated realization via doubling only their respective first grapheme respectively. The other two digraphs (i.e., [tʃ] <ch> and [ʃ] <sh>), however, could better be left to be represented using the current working graphemes since there are no challenges of transfer of skills with such graphemes.

Hence, with regard to script, the Ge'ez script and the Roman based script of Oromo have no relation and this will necessitate the Oromo mother tongue learners to learn Amharic, the federal official working language of the country. In addition to this, despite the fact that Oromo and English share similar Roman based script and some graphemes, the difference in phonetic value of some consonant in both languages negatively affect skill transfer between the two languages and this situation is also true for Hadiyya (Samuel 2017:113).

3.7.5 Maximum ease of reproduction

Maximum ease of reproduction has to do with writing and reproduction of a text. Writing may refer to hand writing, typing on a typewriter, computer, or cellular phone or using a computer to print. A reproduction of a text may be by photocopying, computer printing, any of numerous commercial printing methods, or by publication on the internet (Decker 2014: 42). The design of orthography for a certain language needs the consideration of typing and printing facilities, though this is not the first and most important. From this criteria point of view, the Roman based orthography of Oromo has symbols, all of which exist on the computer keyboard that are compatible and simple to share and read documents on another computer. Thus, from the principles of the maximum reproduction point of view, the Roman based script which Oromo uses is convenient to reproduce since all characters of the orthography are available on a computer keyboard and is compatible share documents.

On the other hand, the issue of the economy is an important that has to be discussed in relation to the orthography of the language. The number of the graphemes and mode of writing words should be in a way that saves space, time and cost. The current Oromo orthography is not economical as compared to other scripts like Ge'ez script. But, transparency is preferred as

compared to the economy. From this point of view, the current orthography of the language has the problem of transparency regarding the representation of the geminated form of the digraphs <dh>, <ph> and <ny> i.e. (/d/, /p/ and /n/) respectively, due care must be taken while marking the geminated form of these digraphs. If they are to be marked, there are two options to mark the germinations of the above digraphs. The first is to double only the first letter of the digraphs as in <hodddhuu> 'to sew' or doubling the digraph as a whole as in <hodhdhuu> 'to sew'. Hence, the second form of marking gemination to maintain transparency is not economical as this will not save space, time and monetary cost. The first option is more economical and solves the problem of transparency.

3.8. Reform of Oromo orthography

The current Roman based orthography has been used to write Oromo for more than twenty five years without modification. As it has been discussed in the above subsections, concerned body, i.e., the standardization committee of Oromo did not make any attempt to modify the orthography of the language despite the existence of orthographic related issues to be considered. In this regard, the research participants confirmed the above issues during FGDs as follows.

The standardization committee of Oromo did not make any reform on the current working orthography of Oromo since 1991. It was Oromia education bureau, for which the right and responsibility is not given by the state government of Oromia, that reformed the order of the alphabets of Oromo. The bureau reformed it without consulting us. Before the reformation by the bureau, the order of the alphabet was 'A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, Ch, Dh, Ny, Ph, Sh' and they reformed it into 'L, A, G, I, M, S, N, U, T, R, E, F, O, D, H, K, J, X, Dh, Q, Y, W, N, B, Ch, Ny, C, P, Sh, V, Ph, Z'

(See Appendix G, FGD 1, 16).

The data suggest that the standardization committee has not made any kind of orthographic reform since the Roman based orthography has been designed to write Oromo in 1991. But, this does not mean that the orthography of Oromo is perfectly designed at its very beginning. There

are unsettled issues of orthography that have been discussed in the previous subsections. It was only education bureau Oromia, for which the right and responsibility is not given by the state government of Oromia, that reformed the order of alphabetic of Oromo for the first time. The bureau consulted nobody when it reformed the order of the alphabet from 'A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, Ch, Dh, Ny, Ph, Sh' into 'L, A, G, I, M, S, N, U, T, R, E, F, O, D, H, K, J, X, Dh, Q, Y, W, N, B, Ch, Ny, C, P, Sh, V, Ph, Z'.

Members of the standardization committee of the language state that the bureau did not make any discussion with us or with any of the higher bodies of the government. Regarding this, one of the research participant states as to how reform on orthography should be made in the FGD. The participant states that if there are scripts and orthographic issues to be considered, it has to be reformed after it has been studied and accepted by both speakers of the language and state government of Oromia (See Appendix G, FGD 1, 25). He further states that nobody or no individual has the right to make any reform without the consultation of the concerned body as it has been decided legally. Hence, if there is a need of language related reform, it has to be legally reformed in consultation with the concerned body of the government. The participant argues that the education bureau of Oromia regional government did contrary to the regulations of the region and if there is alphabet order and orthographic related problems, it has to be supported by research and the result of the study has to be presented to the concerned body. As it has been argued during the discussion, if any reform of Oromo orthography is needed at all, the right of making any reform has been given to the standardization committee of Oromo. The current working alphabet order of the language has been decided to acquire its current status through a legal decision. It is also mentioned during the focus group discussion that changing the order of the alphabet or making a reform to already accepted alphabet order without consulting the speaker, the government is an illegal act (See Appendix G, FGD 1, 25). The participant also noted that the alphabetic order of Oromo has been stated under article five of the constitution of the regional state as 'A, B, C, D, E, F...' not as 'L, A, G, I, M,...' and , hence, changing this on which the speaker lost their life is illegal. Even concerned higher bodies of the regional state government and the speakers have not been aware that the reform has been made and there were strong oppositions from the speakers' side. The right and responsibility of education bureau of

the regional state of Oromia is to teach the language; but is not given by the law to make any reform and the bureau has committed this mistake (See Appendix G, FGD 1, 25).

The criterion behind the reform of the order of the alphabet of the language by Education Bureau of Oromia is the issue to be discussed. The members of the standardization committee of Oromo state in the FGDs as follows.

Education Bureau of Oromia first collected texts and forms a text corpus. Secondly, the frequency with which each character used to write a word has been measured and order of the character has been set. Based on the reform, Education Bureau of Oromia published textbooks. But, the textbooks were not used more than two years as the speakers did not accept the reform made. Even we and the higher regional government bodies were not aware about the reform. Moreover, the reform did not take acceptance of the speakers, transfer of skill and other criterion into consideration. Taking the frequency of the character is not the only criterion for setting the order of the character. Considering its suitability for teaching and learning the language and other issues must be given due care. The right and responsibility of working on the language related issues in general and orthographic related issues in particular has been given to the standardization committee of Oromo legally. Hence, we are the right body to make any language related reforms in consultation with the speakers and the concerned highest body of the state government (See Appendix G, FGD 1, 26).

As the above quote depicts, the orthographic reform was made using a corpus analysis method. First, the Education Bureau of Oromia collected texts from various spheres and developed Oromo text corpus. Then, the frequency of each character was measured and the alphabetic order of the language was set based on the frequency of occurrence. Following this reform, all the textbooks of primary school were changed. But, the textbooks which were developed and published based on the reform were not used more than two years. There was opposition from the speaker side regarding the orthographic reform made by the bureau. Finally, the regional

government of Oromia immediately reversed the reform into its previous alphabet order due to a number of reasons. First, the bureau did not make a discussion with all concerned bodies. Secondly, as the above quote shows, even the higher government bodies of the state were not aware about the reform.

However, still the current official alphabet arrangement itself needs reconsideration with respect to the alignment of monographs with their respective digraph as it is depicted in table 13 below.

Table 13: The current official orthography of Oromo

Upper case	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Lower case	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
Letter name	[a]	[ba]	[tʃˈa]	[da]	[e]	[fa]	[ga]	[ha]	[i]	[dʒa]	[ka]	[la]	[ma]	[na]	[o]	[pa]
Upper case	Q	R	S	T	U	V	W	X	Y	Z	Ch	Dh	Ny	Ph	Sh	
Lower case	q	r	s	t	u	v	w	x	y	z	ch	dh	ny	ph	sh	
Letter name	[kˈa]	[ra]	[sa]	[ta]	[u]	[va]	[wa]	[tˈa]	[ja]	[za]	[tʃa]	[dʒa]	[na]	[pˈa]	[ʃa]	

The current alphabetic order of the language is not systematic; particularly with regard to the order of digraphs in the alphabet. Regarding the order of the alphabet, one of the members of the standardization committee of Oromo states during the FGDs as follows.

Regarding the alignment of Oromo alphabet, it will be a more easier to learn the alphabet of the language if it is rearranged as 'A, B, C, Ch, D, Dh, E, F, G, H, I, J, K, L, M, N, Ny, O, P, Ph, Q, R, S, Sh, T, U, V, W, X, Y, Z'. This will make learning easy especially for beginners. This will help to learn the digraphs easily.

(See appendix II, FGD 2, 22).

The above quote shows that there is a need of rearranging the order of the alphabet at least from some of the members of the standardization committee point of view. The above arrangement of alphabet stated during the focus group discussion helps to learn the digraphs easily. As it has

been noted by the participant, the order will be systematic and easy to learn if the digraphs are placed immediately after their respective monographs. The alphabetic order of Hadiyya, after it has been reformed, look like the recommended alphabet order for Oromo with only few exceptions (see table 14). Still, the present study argues that the alphabet order has to be rearranged. In addition, the frequently occurring glottal stop has not been represented with grapheme and included in the alphabetic order of the language. It is recommended to represent the phoneme with grapheme <ʔ> and placed the last part of the alphabet order. Hence, the present study recommends the following arrangement of the Oromo alphabet.

Table 14: The recommended orthography of Oromo

Upper case	A	B	C'	Ch	D	Dh	E	F	G	H	I	J	K	K'	L	M	N
Lower case	a	b	c'	ch	d	dh	e	f	g	h	i	j	k	k'	l	m	n
Letter name	[a]	[ba]	[c'a]	[tʃ'a]	[da]	[d̥a]	[e]	[fa]	[ga]	[ha]	[i]	[dʒa]	[ka]	[k'a]	[la]	[ma]	[na]
Upper case	Ny	O	P	P'	R	S	Sh	T	T'	U	V	W	Y	Z	Zh	ʔ	
Lower case	ny	o	p	p'	r	s	sh	t	t'	u	v	w	y	z	zh	ʔ	
Letter name	[ɲa]	[o]	[pa]	[p'a]	[ra]	[sa]	[ʃa]	[ta]	[t'a]	[u]	[va]	[wa]	[ja]	[za]	[ʒa]	[ʔa]	

The order of the alphabet of Oromo could be modified as we can see in the above recommended Roman-based Oromo orthography. The digraphs, which are ordered separately after the monographs in table 14, follow now immediately the related monographs. In other words, the arrangement of graphemes in the recommended orthography of Oromo shown in table 14 is systematic which goes from monographs to digraphs and follows similar orthographic shape in the initial graphemes. On the other hand, such type of alphabet order is also true for, for instance, Hadiyya with little exception (Samuel 2017:115). To him, such kind of arrangement of the alphabets is suitable to easily recognize the digraphs at early stages of learning of the language. He further states that such order of the alphabet is important for arranging lexical entry during dictionary compilation. This argument can also work for Oromo. Here, the only difference in the recommended Oromo alphabet order is the representation of ejectives i.e., <c'> /tʃʔ/, <k'> /kʔ/,

<p'> /p'/, <t'> and the glottal stop /ʔ/ with the IPA <ʔ>. The phoneme /s'/ is not recommended to be represented with a grapheme by the present study as it can be replaced by the nearest sound called /t'/ <x>.

3.9. Summary

Oromo has been written using various scripts like the Ge'ez, Arabic and Roman based scripts until the end of 1990. It was in 1991 that Oromo was decided to be written officially using the Roman based script following the change of language policy. Developing orthography is a vital aspect of language standardization. In the same manner, revising and standardizing the existing orthography of a language (if it demands) is almost equally important. From this point of view, the current Roman based orthography of Oromo has been used for about 27 years. The current orthography of Oromo has been described and evaluated in light of different principles.

As it has been discussed in the previous consecutive subsections, <ts> /s'/ unnecessarily introduced into the phonemic inventory of the language as it only occurred in few borrowed words like <tsabala> and this phoneme can be written using the nearest phoneme /t'/ as in <xabala>. The phoneme /z/ which is introduced into the language, but not included into both the phonemic inventory and the alphabet order of the language is being written using various graphemes like <zh>, <zy>, <zn> and <j> as <televizhiinii>, <televizyiinii> and <televijiinii> respectively. On the other hand, the unsettled issue about the graphematic representation of glottal stop or sound represented by /ʔ/ in the IPA inventory is one of orthographic standardization problems to be discussed and normalized. Despite its frequent occurrence in the language and its inclusion as an independent phoneme in the phonemic inventory of the language, the grapheme for this phoneme is not still devised and included in orthography of the language.

There are also limitations regarding the marking of geminated realization of phonemes represented by digraphs in Oromo. Except the diagraphs, all consonants and vowels that have the geminated and length realization are marked by doubling the grapheme that represent the sound and are canonical. To make the orthography of Oromo more transparent, it is recommendable to

mark the geminated realizations of sounds represented by digraphs as the issue of the representation of gemination has been one of the factors for the decision to adopt the Roman based script for writing Oromo. Concerning maximum transfer of skills, representation of the ejectives <x> [tʰ], <q> [kʰ], <c> [tʃʰ] and <ph> [pʰ] has a negative impact on the transfer of reading skills. Therefore, it is better to replace the graphemes of the Oromo ejectives <x>, <q>, <c>, <ph> by the <tʰ>, <kʰ>, <cʰ>, and <pʰ> and marking their geminated realization via doubling the first grapheme respectively (see section 3.7.4).

The alphabetical arrangement of the graphemes of Oromo is also one of orthographic standardization issues that need further consideration. The current alphabetic alignment of characters of the language is not systematic; particularly with regard to the order of digraphs in the alphabet. The digraphs are ordered separately from the monographs. This may have a negative impact on learning. In addition, the frequently occurring glottal stop has not been represented with grapheme and included in the alphabetic order of the language. A better arrangement of an alphabet is suggested for future standardization (see table 14).

CHAPTER FOUR

LEXICAL ELABORATION OF OROMO BY THE COMMITTEE

4.0 Introduction

This part of the study addresses internal and external means of lexical elaboration of Oromo. The standardization committee of Oromo used semantic extension, compounding, abbreviations, blending to internally elaborate Oromo. The committee also applied borrowing and loan translation to externally elaborate the language. Hence, in this chapter, detailed internal and external means of lexical elaboration which are applied by the committee are presented in the sections to come.

4.1 Internal means of lexical elaboration

Internal method of lexical elaboration is one of the elaboration techniques through which any language could be enriched. Oromo uses various internal methods of lexical elaboration techniques to enrich its vocabulary. Semantic extension, compounding, blending and abbreviation are internal means the standardization committee of Oromo used to elaborate the lexical of Oromo.

4.1.1. Semantic transfer method

Semantic transfer is a productive strategy for elaborating the lexicon of Oromo. As to Cabre (1999:93), semantic modification is the major criterion of classifying terms by semantic methods. On the other hand, Akmajian et al (2010:30-31) suggests six ways to describe semantic method. These involve metaphorical extension, broadening, narrowing, semantic drift, reversal and change in part of speech.

A language may use various lexical elaboration techniques to enrich its lexicon. Semantic extension (also referred to as metaphorical extension) is one of the ways in which the meaning of an existing word is extended. Semantic extension is the expansion of vocabulary of one domain to a new domain (Akmajian et al 2010:31). When a language encounters a lexical gap to express a concept, speakers usually take an existing word and extend its meaning to fill the gap. Semantic extension is used in Oromo as one of the methods of lexical elaboration. This method employs language internal resources to expand lexicon of the language. In the following section,

meaning extension in Oromo will be described with illustrative examples. The following data are collected from Wiirtuu volume 1- 8 and Qajeelcha (1998) of the committee to show how meaning extension is used by the standardization committee to elaborate Oromo.

(11) Oromo	meaning 1	meaning 2
a) <i>bo'oo</i>	furrow (of a plough)	order/system/ line of text
<i>damee</i>	parts of a plant	other branch of institutions, bank,
<i>sanyii</i>	seed	ethnic
b) <i>gocha</i>	action	action verb
c) <i>gaalee</i>	shield	phrase
<i>gabatee cubaa</i>	traditional material	key board
d) <i>haantuutee</i>	mouse shaped	mouse of computer
e) <i>fuula</i>	face	page
<i>handhura</i>	navel	center/ core
<i>harka</i>	hand	portion
<i>ija</i>	eye	important
<i>kofa</i>	armpit	angle
<i>onnee</i>	heart	mental acuteness
f) <i>madda</i>	origin	source/ reference

As we can see in the above data, Oromo uses internal means to elaborate its lexicon. Meaning extension is one of such techniques as we can see in (11). Some lexical items in Oromo have multiple meaning in which the base words have homonymic nature. For instance, the word *handhura* 'navel' has multiple layer of meaning in the extended sense (i.e. center /core). Concepts such as 'hand' and 'percent (%)' extended by considering the location of body parts '*harka*' hand'. The lexical meaning extensions presented in (11) are related to farming, materials, and human body parts as in (11a), (11c) and (11e) respectively.

4.1.2 Derivation

Derivation is one of the means through which a language enriches its lexicon. It is a process of creating a new word via adding a prefix, infix or suffix (Refat 2015:20). Derivation is one of the

productive mechanisms in Oromo through which new words are coined for new concepts. Concerning derivation, Oromo is a suffixing language. The following section discusses the derivation means of lexical elaboration in the materials published by the standardization committee of Oromo.

4.1.2.1 Nominalization

The following data are taken from Qajeelcha 2008 to show how derivation is used by the standardization committee of Oromo to elaborate the language. The following examples show how derivation applied by the committee to drive new nouns from bases as it is depicted in (12).

(12) Base	gloss	affix (NMZR)	derived nominals	gloss
a) <i>fakk-at-</i>	'picture-MD'	<i>-noo</i>	<i>fakkannoo</i>	'symbolism'
<i>hidh-at</i>	'tie-MD'	<i>-noo</i>	<i>hidhannoo</i>	'munitions'
b) <i>jig-</i>	'fall'	<i>-suu</i>	<i>jigsuu</i>	'shed'
c) <i>kuma</i>	'thousand'	<i>-ala</i>	<i>kumaala</i>	'major'
d) <i>kurna</i>	'ten'	<i>-ala</i>	<i>kurnaala</i>	'corporal'

As can be seen from (12), most of the above, affixes of Oromo are attached to verbal roots to result in derived nominals as in, for instance, *jig-* 'fall'+ *-suu*→*jigsuu* 'shed'. There are also cases in which same affix is attached to both verbal roots and noun as in, for instance, *hidh-at*- 'tie-MD'+ *-noo*→*hidhannoo* 'munitions' and *fakk-at* 'picture-MD'+*-noo*→*fakkannoo* 'symbolism' respectively.

Moreover, the standardization committee of Oromo applied the affixation of suffixes to verbal roots, nominals and adjectival to designate new concepts. The data in (13) are taken from Qajeelcha (2008) to clearly show the elaboration process.

(13) Base	Gloss	affix (NMZR)	derived nominals	gloss
a) <i>midhag-</i>	'be beautiful'	<i>-ina</i>	<i>midhagina</i>	'aesthetics'
b) <i>nagaa</i>	'peace'	<i>-eenya</i>	<i>nageenya</i>	'security'
<i>dhiyoo</i>	'near'	<i>-eenya</i>	<i>dhiyeenya</i>	'approach'
c) <i>nam-</i>	'person'	<i>-eessa</i>	<i>nameessaa</i>	'personification'
<i>kitaaba-</i>	'book'	<i>-eessa</i>	<i>kitaabeessa</i>	'author'
d) <i>seera</i>	'law'	<i>-ummaa</i>	<i>seerummaa</i>	'legality'
f) <i>dhug-</i>	'be true'	<i>-ooma</i>	<i>dhugooma</i>	'plausibility'
<i>nam-</i>	'man'	<i>-ooma</i>	<i>namooma</i>	'human'

Looking into (13), one can note that new derived nominals are formed via suffixing into verbal, nominal and adjectival bases as in, for instance *midhag*-‘be beautiful’ + *-ina*→*midhagina* ‘aesthetics’, *seera* ‘law’ + *-ummaa*→*seerummaa* ‘legality’, *dhiyoo* ‘near’ + *-eenya*→*dhiyeenya* respectively. Here, we can note that the same nominalizing affix *-eenya* is seen to be attached into a different category of base words as in *dhiyoo* ‘near’ (adjective) and *nagaa* ‘peace’ (noun). Owens (1985:249) states that *-eenya*, *-uma*, *-ina* and *-oma* as in, for instance, *fageenya*, ‘farness’ *obbolyuma* ‘brotherhood’, *diimina* ‘redness’ and *beekoma* ‘knowledge’ respectively to be the abstract formative affixes in the authors description of grammar of Harar dialect of the language. Here, *-ooma* is more productive than *-oma* in Oromo in general. Griefenow-Mewis (2001:25) mentioned another nominalizing affix *-iinsa* as in *bulchiinsa* ‘administration’ *cabiinsa* ‘breakag’

Oromo is being enriched through the affixation of other nominalizing suffixes as seen in (14). The following data are collected from Wiirtuu volume 10 to illustrate some of the formation of derived nominals by the standardization committee of Oromo.

(14) Base	gloss	affix (NMZR)	derived nominals	gloss
a) <i>bar-</i>	'know'	<i>-umsa</i>	<i>barumsa</i>	'education'
b) <i>salph-i-s-</i>	'simplify-ep-CS'	<i>-uu</i>	<i>salphisuu</i>	'simplification'
c) <i>abaar-</i>	'curse'	<i>-sa</i>	<i>abaarsa</i>	'curse, to swear'
d) <i>faal-</i>	'pollute'	<i>-ii</i>	<i>faal-ii</i>	'pollution'

The standardization committee of Oromo is elaborating the language through applying morphological processes. As can be seen from (14), all derived nominal are formed through combining verbal roots and nominalizing suffixes stated under the third column.

Lexicon of Oromo is also enriched through the suffixation of *-soo, -oo, -tee, -taa* and *-lessa* to either verbal roots or to an adjectives. The following data collected from Qajeelcha 2008 illustrate some other examples of the lexical elaboration of the language by the standardization committee of Oromo.

(15) Base	gloss	affix (NMZR)	derived nominal	gloss
a) <i>jig-</i>	'fall'	<i>-oo</i>	<i>jigoo</i>	'land side'
<i>mud-</i>	'encounter'	<i>-oo</i>	<i>mudoo</i>	'defect'
b) <i>milq-</i>	'abscond'	<i>-soo</i>	<i>milqsoo</i>	'controband'
c) <i>bitaa</i>	'left'	<i>-lessa</i>	<i>bitaalessa</i>	'leftist'
d) <i>gal-</i>	'enter'	<i>-tee</i>	<i>galtee</i>	'input'
e) <i>gurgur-</i>	'sell'	<i>-taa</i>	<i>gurgurtaa</i>	'sale'

Bases (roots) of most of the derived nominals stated under the fourth column of data in (15) are verbs. The sample words taken from the codified words by the standardization committee of Oromo, however, show that adjectives are also sources for the formation of new nouns as in, for stance, *bitaa* 'left' + *-lessa* (NMZR) → *bitaalessa* 'leftist'.

One can also note, through looking into the following data in particular and the codified and published words of Oromo in general, that the standardization committee of Oromo enriched the

language via the affixation of *-uu* to the verbal roots as seen in (16). The data in (16) are taken from Wiirtuu volume 10.

16) Base	gloss	affix	verbal noun	gloss
<i>deem-</i>	'go'	<i>-uu</i>	<i>deemuu</i>	'going'
<i>summii</i>	'poison'	<i>-uu</i>	<i>summeesuu</i>	'poisoning'
<i>baay'is-</i>	'multiply'	<i>-uu</i>	<i>baay'isuu</i>	'multiplying'
<i>warshaal-</i>	'assemble'	<i>-uu</i>	<i>warshaaluu</i>	'assembling'
<i>raaj-</i>	'forecast'	<i>-uu</i>	<i>raajuu</i>	'forecasting'
<i>kasaar-</i>	'lose'	<i>-uu</i>	<i>kasaaruu</i>	'losing'

The resultant derived nominal stated under the fourth column are the output of the morphological operation of verbal roots and the nominalizing suffix *-uu* as in *deem*-‘go’ + *uu* (NMZR)→*deemuu* ‘going’. The output of such operation results in a verbal noun as seen in the fourth column of (16). Shimelis (2014:64) states that the verbal noun morpheme *-uu* is transparently directly suffixed to the verbal root or stem. This morpheme is used for the purpose of lexical elaboration in the materials compiled by the committee. Nominalizing morpheme *-a* is also seen to be applied in the processes of lexical elaboration as in, for instance, *arg-i-siis-a* see-ep-CS-NMZR ‘exhibit’. Verbal nouns are extensively used to enrich the language.

The standardization committee of Oromo also used nominalizing suffixes such *-aa* and *-tuu* to designate new concepts. The data in (17) are taken from Qajeelcha (2008:19) to illustrate some of noun formations during the course of lexical elaboration of the language.

17) Base	Gloss	affix	derived nouns	gloss
<i>qor-at-</i>	'study-MD'	<i>-aa</i>	<i>qorataa</i>	'researcher (m)'
<i>qor-at-</i>	'study-MD'	<i>-tuu</i>	<i>qorattuu</i>	'researcher' (f)
<i>ham-at-</i>	'incubate-MD'	<i>-aa</i>	<i>hammataa</i>	'incubator'(m)
<i>dhum-a-t</i>	'consume MD'	<i>-aa</i>	<i>dhumataa</i>	'consumable' (m)

In fact, these are not the only nominalizing affixes in Oromo. There are a number of nominalizing affixes such as *-na* as in *qot-* 'plough' + *-na* (NMZR) → *qot-an* [*qon-na*] plough-(NMZR) 'ploughing'. Here, we can note that the affix *-tuu* can function as both derivational and inflectional affix as in later case this affix can represent gender. Griefenow-Mewis (2001:25) mentions *-a* as a nominalizing affix as in, for instance, *sirba* 'dance'.

4.1.2.2 Verbalization

The main objective of the present study under the above (4.1.2.2) sub topic is not to describe the types, structure (forms) and semantic properties (features) of Oromo causative in general, but to analyze the causative morphemes applied by the standardization committee of Oromo to derive new causative verbs and their causative verbal nouns. What is more important, here, is not to discuss verbalization in Oromo in general, but showing how the standardization committee of Oromo applied verbalization to elaborate the language. Before going into the lexical elaboration via verbalization, disclosing an overview of the existing debate regarding causative morphemes of the language is vital. The types and the numbers of Oromo causative morphemes vary from scholar to scholar and there is no consensus over the issue. For instance, for Gragg (1976:186), the causative morphemes of Oromo are only three; *-s*, *-sis* and *-siis*. Gragg states that *-s* is used for transitivity stative intransitive verbs whereas *-sis* or *-siis* are used for causativizing transitive and intransitive active verbs as in, for instance, *jaam-* 'be blind' > *jaam-s* 'make blind' and *deem-* 'go' > *deem-sis-* 'make go' respectively. To Hayward (1976), *-is* is the underlying causative affix of Oromo and he states the form to be *-s* with elision rule although he did not clearly show the operation. Lloret (1987:145) argues Oromo causative morpheme is only *-s* which occurs cyclically and separated by a morphological epenthetic *-i(i)-*, which is not similar to the phonological epenthetic vowel *i*. On the other hand, to Tolemariam (2009:25), there are five Oromo causative formatives; *-is-*, *-s-*, *-ss-*, *-sis-*, *-eess-* and of these, *-is-* being a common causative marker in the language. Shimelis (2014:101) accepts Gragg's (1976:186) three types of Oromo causative morphemes (*-s-*, *-sis-* and *-siis-*) with the application of Lloret's (1987:145) "vowel length dissimilation rule" in which the vowel length dissimilates to the length of the vowel of the previous syllable for the cases of *-sis-* and *-siis-* (see also Griefenow-Mewis & Tamene 1994:84).

While analyzing the lexical elaboration strategies that are used in the materials published by the authority, the present study identified three kinds of causative morphemes in Oromo that are applied by the standardization committee of Oromo to enrich the lexicon of the language; *-s* [CS¹], *-sis*-[CS²] and *-siis*-[CS³] as we shall see in (18) and (19). Regarding the distribution of the causative morpheme, vowel length dissimilates to the length of the vowel of the previous syllable for the cases of *-sis*- and *-siis*- as it is stated by Lloret (1987:145).

Coming back to the main concern of the present subtopic (i.e., lexical enrichment of the language), apart from the lexical elaboration of Oromo via a nominalization processes, the standardization committee of Oromo applied a verbalization process to designate new concepts. In this regard, it is attested that various causativizing morphemes are applied to derive verbs from stems (forms) to enrich the lexicon of the language as the data collected from various materials of the standardization committee of Oromo depict. The following data are collected from Qajeelcha (2008) to show the operation of verbalization as a means of lexical elaboration of the language by the standardization committee of Oromo.

18) Base & gloss	affix (CS) (VBZR)	causativized verb	Affix (NMZR)	causative verbal nouns	gloss
a) <i>aman</i> - 'believe'	<i>-siis</i> -	<i>aman-siis</i> - believe-CS cause to believe'	<i>-uu</i>	<i>aman-siis-uu</i> believe-CS-NMZR	'causing to 'believe'
b) <i>beek</i> - 'know'	<i>-sis</i> -	<i>beek-sis</i> - know-CS 'cause to know'	<i>-uu</i>	<i>beek-sis-uu</i> know-CS-NMZR	'causing to know'
c) <i>hafars</i> - 'aggravate	<i>-i-s</i> -	<i>hafars-i-s</i> - aggravate-ep-CS 'cause to aggravate'	<i>-uu</i>	<i>hafars-i-s-uu</i> aggravate-ep-CS- NMZR	'causing to aggravate'

<i>d) hor-</i>	<i>-siis-</i>	<i>hor-siis-</i>	<i>-a</i>	<i>hor-siis-a</i>	'husbandry'
'conjugate'		conjugate-CS-		conjugate-CS-	
		cause to conjugate'		NMZR	

As can be seen from (18), the standardization committee of Oromo applied causativization to enrich the lexicon of the language. The bases for all causativization process are verbs as it can be seen under the first column. As seen under the third column, the causativized verbals in turn become a base for the formation of causative verbal nouns through the affixation of nominalizing morphemes *-uu* and *-aa*. The causativizing morphemes *-s-* and *-sis-* extensively applied to elaborate the lexicon of Oromo. This can be proven via looking into the codified words by the standardization committee of Oromo in general and the data in (18) in particular.

There are also a number of other causative morphemes which are evident in various texts of the standardization committee of Oromo and that are applied in the processes of lexical elaboration of the language. The following data are collected from Wiirtuu volume 10 and from Qajeelcha (2008) to illustrate the application of other causative morphemes in the course of elaboration of Oromo.

19)Base& Gloss	affix (CS)	causativized verb	affix	causative verbal nouns	gloss
a) <i>arg</i> - 'see'	- <i>i-siis</i> -	<i>arg-i-siis</i> - see- ep-CS- 'cause somebody to make somebody else see'	- <i>a</i>	<i>arg-i-siis-a</i> see-ep-CS- NMZR	'exhibit'
b) <i>baq</i> - 'melt'	- <i>s</i>	<i>baq-s</i> - melt-CS- 'cause to melt'	- <i>uu</i>	<i>baq-s-uu</i> melt-CS- NMZR	'melting'
c) <i>cab</i> - 'break'	- <i>siis</i> -	<i>cab-siis</i> - break-CS-CS- 'make somebody break something'	- <i>uu</i>	<i>cab-siis-uu</i> break-CS- NMZR	'causing to break'
d) <i>dubb</i> - 'talk'	- <i>i-s</i>	<i>dubb-i-s</i> - talk-ep-CS- 'cause to talk or great'	- <i>uu</i>	<i>dubb-i-s-uu</i> talk-ep-CS- NMZR	'make to greet'
e) <i>qab</i> - 'hold'	- <i>siis</i> -	<i>qab-siis</i> - 'hold-CS- 'make to hold'	- <i>uu</i>	<i>qab-siis-uu</i> hold-CS— NMZR	'make to hold'

There are other causativising morphemes that are used to elaborate the lexicon of Oromo apart from the morphemes we have seen in (18). As can be seen from (18) and (19) the causativizing morphemes *-sis-* and *-siis-* are applied when a stem ends with closed syllable (CVVC) and short vowels (CVCC or CVC) respectively. On the other hand, structurally, *-s* and *-is* are similar as *-i-* of *-is-* is considered an epenthetic vowel.

Apart from the application of causativization in the process of lexical elaboration of Oromo, there are other means of verbalization. One such method is the suffixation of *-at*, *-moo* as in, for instance, *furdāt* 'get fat and 'duroom 'get rich' respectively. Owsn (1985:248) states that such suffixes can be used to form inchoatives. The passive *-am* are also applied to elaborate the language in the materials compiled by the committee as in, for instance, *filatam* (*fil-at-am*

choose-MD-PS) ‘be chosen’. This formative is also attested in the course of inreachment of the language (see also Ali & Zaborsk (1990:53). Converb formatives are also seen in the lexical elaboration of the language as in, for instance, *hojj-at-Ø-ee* (work-MD-3ms-CPL).

4.1.2.3. Adjectivization

The objective of this subtopic is not to describe adjectivization in Oromo in general, but to examine how adjectivization is used by the standardization committee of Oromo to enrich the lexicon of the language. Adjectivization is also one of the lexical elaboration techniques applied by the standardization committee of Oromo. The following data are collected from Qajeelcha (2008) as an illustration of derivation of new adjectives from existing verbs.

20) Base	gloss	affix	derived adjective	gloss
<i>rincic-</i>	'conserve'	<i>-aa</i>	<i>rincicaa</i>	'conservative (m)'
<i>dheegal-</i>	'be deaf'	<i>-aa</i>	<i>dheegalaa</i>	'deaf (m)'
<i>dheegal-</i>	'be deaf'	<i>-tuu</i>	<i>dheegaltuu</i>	'deaf (f)'

As seen in (20), the standardization committee of Oromo codified the above new adjectives after operation of verbal roots with adjectivizing affixes such as *-aa* and *tuu* for masculine and feminine respectively. Adjectivization is not as productive as nominalization and verbalization in the process of lexical elaboration of the language.

4.1.3. Compounding

Compounding is the way which combines at least two words into one word according to certain order Cheng 2018:251. The elements can be characterized as words, stems, or lexemes, depending on the language and the elements that can be combined can be categories of noun, verb, adjective or adposition. Words created by compounding are called compounds which are a new form consisting of two or more constituents but reflect a single concept. Compounds can be divided into different types. According to word class, it can be divided into compound nouns, compound verbs, and compound adjectives. Among them, compound nouns have the largest proportion (Cheng 2018:251). According to internal structure, compounds can be divided into noun + noun, noun + verb, verb + verb. A compound can also be classified into endocentric and

exocentric based on the relation between semantic and lexical category of the constituents of the compound and the output compound. A compound that does not have a (semantic) head; or a compound whose head falls outside of the construction is exocentric whereas an endocentric compound has a semantically hyponymical relationship between the head and the compound Benczes (2015:54). In the case of appositional compounds, the meaning of a compound results from the equal contribution of its constituents and each element is a head in its own right. Compounds can be divided into different types. Baed on word class, it can be divided into compound nouns, compound verbs, and compound adjectives. Hence, from the above conceptual lexical elaboration point of view, compounding in Oromo will be described in the following subsections.

4.1.3.1 Lexical elaboration via compounding

The process of compound, phrasal lexeme and proper phrase distinction is often an area of debate in the discipline of linguistics. Masin (2009:254) states that in the constructionist framework there is no clear-cut division between lexical items and syntactic structures: the construction – intended as a conventionalized association of a form and a meaning or function – is the basic unit of linguistic analysis; and the syntax-lexicon distinction is rather a matter of degree (see also Goldberg (1995, 2006). This is due to the fact that compounds and syntactic structures resemble each other in their hierarchical and recursive structures (Spencer 1991: 310). There is almost no study regarding the distinction of Oromo compounds and phrases. Some criteria of distinguishing Oromo compounds from other construction are discussed under section 4.1.2.4 though some criteria may not be valid for all compounds of the language. The discussion is not exhaustive as the main objective of the present study is not to make a distinction between Oromo compounds with other construction. Rather, it is aimed at disclosing some of compounding strategies that are applied to enrich the lexicon in the text corpus this study consulted. Hence, this study residue such area of study for further detaile evestigation.

Oromo compounds can be classified into verbal, nominal and adjectival compounds (Temesgen 1993:34). Compounds in Oromo are formed by the combination of various lexical categories and yield compounds of various categories. The objective of this subtopic is not, basically, to describe

the compounds of Oromo in general, but to analyze how compounding is used for lexical elaboration by the standardization committee of Oromo. The compounding strategies that are used to enrich the lexicon of Oromo are discussed in the following subsections.

4.1.3.1.1 Nominal compounds in Oromo

Two nouns can combine to form various kinds of compound nouns in Oromo. Nominal compounding is very productive in Oromo word formation. This section discusses N+N→N compounding of Oromo. The following data are taken from Wiirtuu volume 9-10 of the standardization committee of Oromo.

(21) Noun	noun	compound noun	gloss
<i>bu'ura</i> 'base'	<i>misooma</i> 'development'	<i>bu'uramisoomaa</i>	'infrastructure'
<i>caffee</i> 'council'	<i>tasgabbii</i> 'security'	<i>caffee-tasgabbii</i>	'security council'
<i>gibira</i> 'tax'	<i>galii</i> 'income'	<i>gibiragalii</i>	'income tax'
<i>gurmuu</i> 'group'	<i>hojii</i> 'task'	<i>gurmuuhojii</i>	'task force'
<i>jaalala</i> 'love, loyalty'	<i>afaan</i> 'language'	<i>jaalalaafaanii</i>	'language loyalty'
<i>kubbaa</i> 'ball'	<i>saaphana</i> 'net'	<i>kubbaasaaphanaa</i>	'volleyball'
<i>mana</i> 'house'	<i>barumsa</i> 'education'	<i>manabarumsaa</i>	'school'
<i>mana</i> 'house'	<i>citaa</i> 'grass'	<i>manacitaa</i>	'house made of grass'
<i>mirga</i> 'right'	<i>kitaabessa</i> 'author'	<i>mirgakitaabessaa</i>	'copy right'
<i>qabeenya</i> 'wealth'	<i>gamtaa</i> 'union'	<i>qabeenya gamtaa</i>	'common wealth'
<i>siree</i> 'bed'	<i>muka</i> 'wood'	<i>sireemukaa</i>	'bed of wood'
<i>sirna</i> 'system'	<i>barumsa</i> 'education'	<i>sirnabarumsaa</i>	'curriculum'

The compounds of Oromo shown in (21) can be discussed from the point of view of compound nouns' head and modifier. In the N + N compounds, the role of the second noun is to make the meaning of the first noun more specific (see also Temesgen 1993:34). In compound words such as, *sirna* 'system' *barumsa* 'education'→*sirna barumsaa* 'curriculum', *siree muka* (bed + of wood) 'bed of wood', *mirga kitaabessaa* (right + of author) 'copy right', *sirna barumsaa* (system + of education) 'curriculum' *gurmuu hojii* (group + task of) 'task force' etc. the first constituents are heads of the compounds. More specifically, in line with the arrangement of modifier and

head of compounds, Oromo is left headed compounds as the above data shows. From the semantics point of view, all the nominal compounds stated in (21) endocentric compounds as the meaning of each compound emanates from the meanings of each constituent of the respective compounds. In case of N + N construction of compounds where compounds are transparently analysable and thus need not be memorized (or listed) and that N + N compounds (e.g., dog house) are equivalent to possessive + N phrases (dog's house) (Shimelis 2014:204; Bauer 2003: 135f.). The compounds stated under data (21) can be clear examples of such compound constructions. Such compounding lays in-between syntax and morphology.

As can be noted from (21), from the formal (structural) point of view, it is difficult to clearly differentiate Oromo compounds from genitive phrase construction particularly when the elements of compounds are both nouns (see Shimelis 2014:211). Shimelis (2014:212) also states that although the construction is formally a genitive phrase, it is semantically used as a lexicalized compound in the meaning representing a specific concept. There are also a number of features (including the semantic criteria) to distinguish compound construction from the genitive phrases in Oromo (see section 4.1.2.4).

This is not the only way of compound formation. The following data taken from Qajeelcha 2008 and from the standardization committee of Oromo Wiirtuu volume 10 2005) shows compounding through deletion of elements.

(22) Noun + noun	nominal compounds	gloss
<i>boca</i> 'shape' + <i>onnee</i> 'heart' →	<i>boconnee</i>	'heart shaped'
<i>caasaa</i> 'structure' + <i>luga</i> 'language' →	<i>caasluga</i>	'grammar'
<i>galmee</i> 'record' + <i>argii</i> 'which can be observed' →	<i>galmargii</i>	'documentary film'
<i>goda</i> 'cave' + <i>ambaa</i> 'remaining' →	<i>godambaa</i>	'museum'
<i>mana</i> 'house' + <i>kuusa</i> 'ware' →	<i>mankuusa</i>	'ware house'
<i>mana</i> 'house' + <i>seera</i> 'law' →	<i>manseera</i>	'court'
<i>mana</i> 'house' + <i>soora</i> 'food' →	<i>mansooraa</i>	'restaurant'
<i>qaama</i> 'body' + <i>jiraa</i> 'real, exist' →	<i>qaamjiraa</i>	'real mass'
<i>roga</i> 'side' + <i>kofa</i> 'angle' →	<i>rogkofa</i>	'trigonometry'

As we can see in (22), the second compounding strategy is the deletion of final vowels of the first constituent and taking the second constituent as it is. This compounding method is attested in all of the above data with the exception of the last three compounds. In compound words such as *manseera* (*mana* 'house' + *seera* 'law') 'court', *mankuusa* (*mana* 'house' + *kuusa* 'ware') 'house ware', *boconnee* (*boca* 'shape'+ *onnee* 'heart') 'heart shaped'etc., the first constituents are the heads of the compounds. This strategy is very productive in endocentric compounds as well.

Oromo uses some lexical items frequently as one of the constituents of compounds to designate new concept. Accordingly, in Oromo compounding, there are some Oromo words which are employed as sources for the coinage of scientific and technical terms. For instance, various fields of study or area of knowledge is designated by employing the Oromo word *xin-*'-logy' from Oromo word *xinxala* 'study' that is equivalent to the English morpheme '-logy' which came from the Greek term '*logus*'. The following data taken from Qajeelcha, materials compiled by standardization committee of Oromo, demonstrate how the committee used this word for lexical elaboration of the language as it is depicted under (23).

(23) <i>xin-</i> + Oromo word	compound	gloss
<i>xinxala</i> 'study' + <i>naannoo</i> 'environment'	<i>xinnaannoo</i>	'ecology'
<i>xinxala</i> 'study' + <i>sammuu</i> 'brain'	<i>xinsammuu</i>	'psychology'
<i>xinxala</i> 'study' + <i>qooqa</i> 'language'	<i>xinqooqa</i>	'linguistics'
<i>xinxala</i> 'study' + <i>latii</i> 'morpheme'	<i>xinlatii</i>	'morphology'
<i>xinxala</i> 'study' + <i>lubaawa</i> 'living things'	<i>xinlubbaawa</i>	'biology'
<i>xinxala</i> 'study' + <i>dhibee</i> 'disease'	<i>xindhibee</i>	'pathology'

As the above data shows, *xin-* (-logy) is very productive to designate a number of terms in Oromo. Such lexical elaboration technique is frequently used by the Oromo standardization committee during the codification process of the language. Almost all compounds are formed via the combination of the first and the second constituents through deleting CVCV of (the last segments) the first constituent. Hence, Oromo uses the term *xin-* as first element of a compound to designate a number of fields of specialization.

There are also adposition + derived noun combinations that result in nominal compounds as seen in (24). The data are collected from the standardization committee of Oromo Wiirtuu volume 10 2005.

(24)	post-positions	derived noun	nominal compound	gloss
a)	<i>dura</i> 'front'	+ <i>taa'a</i> 'sitting' →	<i>dura taa'a</i>	'chairman'
	<i>dura</i> 'before'	+ <i>dhaloota</i> 'birth' →	<i>dhaloota dura</i>	'BC'
	<i>irra</i> 'on'	+ <i>antoo</i> 'becoming above' →	<i>irrantoo</i>	'numerator'
	<i>keessa</i> 'in'	+ <i>deebii</i> 'returning' →	<i>keessa deebii</i>	'revision'
(b)	<i>gidduu</i> 'between'	+ <i>galeessa</i> 'entering' →	<i>giddugaleessa</i>	'average'
(c)	<i>irra</i> 'on'	+ <i>fuudha</i> 'to marry' →	<i>irfuudha</i>	'bigamy'
	<i>irra</i> 'on'	+ <i>dubbannoo</i> 'to talk' →	<i>irdubbannoo</i>	'negotiation'

All the words under the first column are post-positions (Baye 1985:119 & Griefenow-Mews 2001:50). As it is depicted in (24), compound noun can also be formed via the combination of post-positions and nouns. Under (24a) there, is no deletion of any parts of both constituents during compound formation. To the contrary, vowel and CV deletions are made under (24b) and (24c) respectively in the processes of lexical elaboration. Apart from this, compound noun can be formed through the combination of verb + noun as in *qotee bulaa* 'farmer' (*qotee* 'plough' + *bulaa* 'he who live'), *horsisee bulaa* 'he who live by breeding' (*horsisee* 'he who breed' + *bulaa* 'he who live'). But such compounds are not many.

There is also another possible combination of nominal compounds in Oromo which are formed by juxtaposed constituents i.e. noun + adjective. The following compounds are taken from the standardization committee of Oromo Wiirtuu volume 5 2005 to analyze the lexical elaboration techniques used by the standardization committee of Oromo.

(25) Noun	adjective	compound noun	gloss
<i>asiidii</i> 'acid' + <i>goggogaa</i> 'dry' →		<i>asiidii goggogaa</i>	'anhydride acid'
<i>dhaadannoo</i> 'slogan' + <i>ijoo</i> 'main' →		<i>dhaadannoo ijoo</i>	'main slogan'
<i>dogoggora</i> 'error' + <i>guutalaa</i> 'absolute' →		<i>dogoggora guutalaa</i>	'absolute error'
<i>gabaa</i> 'market' + <i>bilisaa</i> 'free' →		<i>gabaa bilisaa</i>	'free market'
<i>hiika</i> 'meaning' + <i>kallattii</i> 'literal' →		<i>hiika kallattii</i>	'literal meaning'
<i>hiika</i> 'meaning' + <i>galchaa</i> 'contextual' →		<i>hiika galchaa</i>	'contextual meaning'
<i>hogganaa</i> 'boss' + <i>eeggataa</i> 'acting' →		<i>hogganaa eeggataa</i>	'acting boss'
<i>humna</i> 'energy' + <i>kuufamaa</i> 'potential' →		<i>humna kuufamaa</i>	'potential energy'
<i>iddattoo</i> 'sample' + <i>kaayawaa</i> 'purposive' →		<i>iddattoo kaayawaa</i>	'purposive sampling'
<i>iyyaattoo</i> 'badge' + <i>haphomaa</i> 'lable' →		<i>iyyaattoo haphomaa</i>	'adhesive badge'
<i>murtii</i> 'decision' + <i>mayyeessa</i> 'last' →		<i>murtii mayyeessa</i>	'last decision'
<i>sanbata</i> 'sabbath' + <i>guddaa</i> 'big' →		<i>sanbata guddaa</i>	'Sunday'

The right-most adjectives modify the left-hand head nouns. The above noun + adjective nominal compounds are semantically endocentric ones. The adjectives on the right side modify different features of the head nouns by specifying the size (e.g. *guddaa* 'big'), situations and feelings (e.g. *suukanneessaa* 'grave'), color (e.g. *adii* 'white') etc.

4.1.3.1.2 Verbal compounds

Verbal compounds in Oromo are not frequently occurring. Verbal compounds in Oromo are formed by the combination of ideophone and verb. For instance, *cal jedhe* 'kept silent' (*cal* 'silent' + *jedhe* 'said'), *gad jedhe* 'tilt down' (*gad* 'down' + *jedhe* 'said') are some of such verbal compounds in the language. Such combination of Oromo ideophones and the verb *godhe* 'made' can result in verbal compounds. The possible combinations are given in (26) below. The data collected from Wiirtuu volume 10 2005 show some of the combinations applied in the courses of creating verbal compounds.

(26) Ideophone	gloss	verb	verbal compound	gloss
<i>biliq</i>	'kill'	<i>godhe</i> 'made'	<i>bilq godhe</i>	'he killed'
<i>cal</i>	'silent'	<i>godhe</i> 'made'	<i>cal godhe</i>	'he made silent'
<i>gad</i>	'down'	<i>jedhe</i> 'said'	<i>gad jedhee</i>	'he stooped down'
<i>qacam</i>	'caught'	<i>godhe</i> 'made'	<i>qacam godhe</i>	'he caught (suddenly)'

As can be seen from (26), most of the verbal compounds that are formed via the combination of ideophones and the verb *godhe* 'made' are semantically endocentric

4.1.3.1.3. Adpositional compound

Another type of compound in Oromo is compound adpositionals. Such types of compounds are formed through the combination of adpositions as the following data (27) collected from Qajeelcha 2008 depicts.

(27) adposition + adposition	adpositional compound	gloss
<i>dura</i> 'front' + <i>duuba</i> 'back' →	<i>duraa duuba</i>	'back and front'
<i>gara</i> 'to' + <i>fulduraa</i> 'front' →	<i>garaful duraa</i>	'towards front'
<i>irra</i> 'on' + <i>jala</i> 'under' →	<i>irraa jala</i>	'top to bottom'

As the above data shows, all most all of the compounds illustrate directions in relation to pre- and post-positional. But, adpositional compounds are not very common like other types of compounds in Oromo.

4.1.3.1.4 Compound adjective

Compound adjectives are not productive means of lexical elaboration in Oromo. It cannot be, however, concluded that there is no instance of such word formation. The following data collected from Qajeelcha (1998) show some of the combinations that result in compound adjectives.

(28) Noun /adjective + noun /adjective	compound adjectives	gloss
<i>amala</i> 'conduct' + <i>qabeessa</i> 'that possess'	<i>amala-qabeessa</i>	'good conduct'
<i>garaa</i> 'stomach' + <i>jabeessa</i> 'that possess'	<i>gara-jabeessa</i>	'heartles, hard-hearted'
<i>gowwaa</i> 'fulish' + <i>namaa</i> 'of man'	<i>gowaanamaa</i>	'foolish'
<i>ija</i> 'eye' + <i>jaamaa</i> 'blind'	<i>ijajaamaa</i>	'blind'
<i>nama</i> 'human' + <i>nyaataa</i> 'that eat'	<i>namanyaataa</i>	'killing'

As can be seen in (28), adjectival compounds in Oromo can be formed via the combination of noun and noun as well as noun and adjectives. Compound adjectives can also be formed via adjective and noun as in, for instance, *gowwaa* 'fulish' + '*namaa* 'of man'→*gowwaa namaa* 'foolish'. This lexical elaboration technique is not extensively used by the standardization committee of Oromo as it is rare in the language (see section 5.2.1.1.4.2.2 & Temesgen 1993:38).

4.1.3.2 Summary of compound types in Oromo

Compound types are discussed under the previous subsections. Accordingly, there are different combinations of word categories that mainly result in nominal compounds followed by adjectival, adpositional and verbal compounds. Oromo compounds that are discussed under the previous subsections are summarized in the following table in relation to various possible combinations of word categories.

Table 15: Inventory of compound types of Oromo based on word categories

		Right constituent			
Left constituent		Noun	verb	adjective	Adposition
	Noun	<i>gibira galii</i> 'income tax'	-	<i>gabaa bilisaa</i> 'free market'	-
	Verb	<i>qotee bulaa</i> 'farmer'	-	-	-
	Adjective	<i>gowwaa namaa</i> 'foolish'	-	-	-
	Adposition	<i>dura taa'aa</i> 'chairman'	-	-	<i>olii gadii</i> 'up and down'
	Ideophone	-	<i>cal jedhe</i> 'he become silent'	-	-

The combination of adposition + adposition, adjective + noun and verb + noun are very rare or not productive as compared to other word classes. The combination between N+N and N+A is a productive combination of word forms in Oromo. In Table 15, *oli gadi* (*oli+ gadi*) (up+down) 'up and down' has the combination of ADP+ADP which results in the adpositional compound.

4.1.2.3 The spelling aspect of compounds

Oromo compounds can also be discussed from the spelling point of view. Oromo compounds are usually being written in two forms. The first and the most common way of writing Oromo compounds is writing them as two separate lexical items with space between the constituents as *gibira galii* 'income tax', *sanbata guddaa* 'sunday', *gurmuu hojii* 'task force' etc. Such kinds of forming compounds are depicted under (21), (25) and (26). The second possible way of writing a compound, as it is attested in the materials this study consulted, is writing compounds as a single lexical item without any space or hyphen as *xinnaannoo* 'ecology' and *xinsammuu* 'psychology'. All of the data stated under (23) can be examples of compounds that have the form of single lexical items which are made of two independent words. Most of the compounds under

(22) formed via deleting some parts of the first constituent and attaching it into the second constituent without deletion of any parts of the second constituent. The third possible ways of writing compounds in Oromo that is evident in the text corpus this consulted is to hyphanate the constituents as in, for instance, *mata-duree* 'title'

Orthographic rule of writing compound words is not clearly stated in Oromo. As a result of the lack of clearly stated orthographic rule (conventions) for writing compounds, there are some inconsistencies observed in writing Oromo compounds. The following compounds are some of illustration for the assertion that there exist inconsistently written compounds even in the materials published by the standardization committee of Oromo itself. The following data are collected from Wiirtuu volume 7-11.

Table 16: Inconsistently written Oromo compounds due to lack of codification

Inconsistently written forms	proposed form	codified by SCO	gloss
a) <i>fiixaa sagalee, fixsaga</i>	<i>mussaga</i>	<i>mussaga</i>	'allophone'
b) <i>gad-aantummaa, gadaantummaa</i>	<i>gadaantummaa</i>	-	'inferiority'
c) <i>hog-barruu, hogbarruu, ogbaruu</i>	<i>ogbaruu</i>	-	'literature'
d) <i>lubbu qabeeyyii, lubbu-qabeeyyii</i> <i>lubbuqabeeyyii</i>	<i>lubbuqabeeyyii</i>	-	'living things'
e) <i>mataduree, mata-duree, mataduree</i>	<i>mataduree</i>	-	'title'
f) <i>nam-tolchee, nama tolchee, namtolchee</i>	<i>namtolchee</i>	-	'man made'
g) <i>og-fakkii, ogfakkii</i> <i>og-barruu, ogbarruu, ogbarruu</i>	<i>ogfakkii</i>	<i>ogfakkii</i>	'drawing'
h) <i>sab-qunnamtii, subqunnamtii</i>	<i>subqunnamtii</i>	-	'mass media'
i) <i>yaadiddama, yaad-hiddama</i>	<i>yaaxxina</i>	<i>yaaxxina</i>	'theory'

As we can see in the above table, the compounds listed under the first column depict inconsistently written forms in different materials this study consulted whereas most of those under the second column are the proposed rule-based forms. Some of these rule are deletion of the final vowel of the first constituent as in *mataa* 'head' + *duree* 'first' → *mataduree* 'title', *gadi* 'down' + *aantummaa* 'next' → *gadaantummaa* 'inferiority' etc. Some of the compounds under the third column are compounds that are codified by the standardization committee of the language

and those that are not yet codified are not mentioned under this column. Compound words stated under the first column are inconsistent with the rule of compound words and they are not economical as compared to the suggested forms by the present study in column two. This is because most of those inconsistently written compound words are formed through vowel deletion, hyphenated and through leaving a space between the constituents. Such inconsistencies will affect the standardization of the language. Hence, the standardization committee of Oromo has to codify such compounds.

4.1.2.4 Defining features of Oromo compounds

The distinction between compounds and multiword-expressions (phrasal words) and in some of the cases even phrases is problematic. As stated by Baldwin and Kim (2010), the question of what constitutes a word is surprisingly complex, and one reason for this is the predominance in everyday language of elements known as multiword expressions (MWEs) and they consist of several words (in the conventionally understood sense) but behave as single words to some extent. For instance, the expression like *by and large*, which any English speaker knows can have roughly equivalent meaning and syntactic function to *mostly*, an adverb. The concept MWE is very complex. Among the problematic characteristic of these expressions are (1) syntactic anomaly of the part-of-speech (POS) sequence preposition + conjunction + adjective, (2) non-compositionality: semantics of the whole that is unrelated to the individual pieces, (3) non-substitutability of synonym words. Baldwin & Kim (2010:269) state that multiword expressions (MWEs) are lexical items that: (a) can be decomposed into multiple lexemes; and (b) display lexical, syntactic, semantic, pragmatic and/or statistical idiomaticity. Giegerich (2005:45) states that in general, compounds may be the product of both productive rules and lexicalization processes, whereas phrases are fully productive; also, compounds may be both semantically transparent and idiomatic, whereas phrases are generally transparent; finally, parts of compounds cannot undergo syntactic operations. The author further states a clear-cut demarcation of the phenomena remains problematic. Indeed, some word combinations display a hybrid behavior and are to be regarded as entities in-between nouns (words) and noun phrases. This is the case of so-called ‘associative ADJ-N combinations’ such as *urban policeman*, *tropical fish*, *medical building*. What is at stake regarding the lexicon-syntax divide is the distinction between

compound and phrase – essentially the distinction between the lexical category N and the syntactic category NP. There are compounds whose form fluctuates between a single-word and a two-word spelling. These are phrasal structures with an increasingly tighter compound status, which forms a continuum from phrases to completely settled compounds and intermediate hybrid formations.

Hence, it is not easy to give a universally accepted definition for compounding and compoundhood differs from language to language. Therefore, proposing a cross-linguistic definition to compoundhood could be regarded as an area still worthy of further investigation (Mitib Altakhaineh 2016:59). The challenges of identifying compounds and phrases of a language also emanate from variation of definition of compound and phrase. What constitutes a compound and a phrase in a language during their formation is also an area that needs further study.

In some language, for instance, in Oromo, it is challenging to structurally classify a compound and genitive constructions. Though the distinction among compounds, multoword expressions and phrase needs further investigation, there are some criteria that distinguish compounds from other constructions. Characteristics of Oromo compounds can be noted via looking the orthographic, morphological, phonological, syntactic and semantic features. Such criteria may not be valid for all compounds in Oromo.

Orthography can be used as a possible criterion to test compoundhood in some languages like the case of Czech and Slovak in which because all compounds are spelled as one word, whereas syntactic phrases are spelled as separate words (Lieber & Štekauer 2009:7). This criterion, however, may not work for other languages. For instance, Oromo compounds are spelled open, hyphenated and closed. Hence, orthographic criterion helps to distinguish Oromo compounds from phrases works only for closed compounds as phrases in Oromo are spelled as separate words as in, for instance, *caasaa* 'structure'+ *cimaa* 'strong'→*caasaa cimaa* 'strong structure'(a phrase) while the compound *caasaa* 'structure'+ *luga* 'language' →*caasluga* 'grammar' is a closed compound. Thus, it can be argued that orthography in Oromo is unreliable, as there is no

consistency in the orthographic representation of compound in the language and a single compounds can be written open, closed and hyphenated as in *nam tolchee*, *namtolchee* and *nam-tolchee* 'man made' respectively.

Compositionality can also be used as distinguishing criteria between compounds and phrases. A multiword is non compositionality if its meaning cannot be deduced from its constituent parts and a compositional if its meaning can be expressed in terms of the meaning of its constituents. Borer (2009:205) shows that compositionality in Hebrew is a reliable criterion to distinguish between compounds, on the one hand and various phrase types, on the other. Here, we can note that being reliable in one language and unreliable in another; surely, does not make a certain criterion valid cross-linguistically (Mitib Altakhaineh 2016:67). In this regard, compositionality can be used as a criterion to distinguish compounds from other constructions such as from phrases and idioms in Oromo. Accordingly, except appositional compounds in which the meaning of the compound is purely compositional as the elements are conjoined without dependency between them (see section 4.1.3.1.3), Oromo compounds are non-compositionals as in, for instance, *gurmuu hojii* 'task force' of meaning which is not transparently related to either a *gurmuu* 'team, group' or *hojii* 'work, task'. Hence, when the compounds of Oromo are noted from the perspective of semantics, the meanings of compounds are not transparently related to either the meaning of their constituents (Shimelis 2014:204). This compound represents a single concept.

Referentiality is another criterion that can be used to distinguish compounds from other constructions. As it is noted by Saeed (2003: 12), referentiality is the relationship by which language hooks onto the world. For instance, the English compound *food lover* the first element of a compound is normally non-referential. Any referential element used to modify a compound in English usually modifies the right element or the head as opposed to the first element or the non-head as in *these* in *these bank accounts* (Mitib Altakhaineh 2016:69). In Oromo too, any referential element to modify a compound modifies the head as in, for instance, *kun* in *kun sirna barumsa* 'curriculum' 'this is curriculum'. Here, *kun* modifies the head *sirna* 'system' and not *barumsa* 'education'.

Lexical collocation is one of the criteria that can be used to draw a borderline between compounds and phrase. Collocations can be broadly defined as recurrent lexical elements which contribute to the text cohesion collocations are also called “restricted lexical collocations” (Tutin 2008:1). Constituents of compounds of Oromo shows an association as, in for instance, *gibira* 'tax' + *galii* 'income of'→*gibira galii* 'income tax'. Here, we can note that the base *gibira* 'tax' is the prominent element while the collocate *galii* 'income' depends on the base.

Concerning the morphological feature, a compound of Oromo can behave just like any simplex word of Oromo can do. For instance, the compound *hoji dhabaa* 'unemployment' can be transformed to abstract compound noun via the adding of the morpheme *-ummaa* as in '*hoji dhabdummaa* (job lost) 'unemployment'. This criterion, however, may not work for all Oromo compounds.

Lexical integrity can also be a criterion to distinguish Oromo compounds from other constructions in that it cannot be interrupted by other elements. In compound construction, the constituents of compound cannot be interrupted with, for instance, affixes. For example, a definite marker of Oromo- *icha* is attached not to the first constituent of a compound; but it is attached to the second constituent of the compounds as seen in (29). The data are collected from the standardization committee of Oromo Wiirtuu volume 10 2005.

(29) compound word and definite marker	coumpound	gloss
a) <i>gibira galii-cha</i> taxt income-DF	<i>gibiragalichaa</i>	'the income tax'
* <i>gibira galii-icha</i> tax income-DF* 'the tax income'	<i>gibiragalichaa</i>	
b) <i>gurmuu hojii-icha</i> group task-DF	<i>gurmuu hojiichaa</i>	'the task force'
* <i>gurmuu-icha hojii</i> group- DF task	<i>gurmichaa hojii</i>	*'the force task'

As can be seen in (29), due to the more acceptability of the attachment of definite marker *-icha* to the second constituent, the non asterisk (*) examples are acceptable compound form. This, however, does not mean that such rules are an absolute.

Apart from the attachment of the derivational and definite marker, inflectional morphemes are most often attached to the modifier of compounds. As a result of just like lexical integrity observed see in any simplex words, Oromo compounds exhibit the disallowing (discouragement) of the insertion of any kind of morphemes in between the constituents of a compound.

The following data are taken from *Wiirtuu* Volume 10 2005.

(30) Compound + inflectional morphemes		gloss
a) <i>manaeen kitaabaa</i>	<i>manneen kitaabaa</i>	
<i>mana-eeen kitaabaa</i>	<i>manneen kitaabaa</i>	
house-PL book	houses of book	'libraries'
b) <i>mana kitaaoota</i>	<i>mana kitaaboota</i>	
<i>mana kitaa-oota</i>	<i>mana kitaaboota</i>	
house book-PL	house of books	'house books'
c)* <i>gibirawwan galii</i>	<i>gibirawwan galii</i>	
<i>gibira-wwan galii</i>	<i>gibirawwan galii</i>	
tax-PL income	taxes income	'income tax'
d) <i>gibira galiwwanii</i>	<i>gibira galiwwanii</i>	
<i>gibira galii-wwanii</i>	<i>gibira galiwwanii</i>	
taxincome-PL tax	incomes	income tax

Here, we can note that, in most of the cases, the non-heads of the compounds bear the inflectional morpheme, not the head. For instance, in (30d), the non-head (modifier) of the compound *gibira galiwwanii* is *galii* 'income' and the head is *gibira* 'tax'. Hence, *galii* takes the inflectional morpheme (plural marker) *-wwan*. There are, however, exceptions as in the case of (30a) in which head takes the inflectional morphemes (Shimelis 2014:222).

Syntactic construction can also be used to test whether compounds of Oromo can have the characteristic features of a simplex word or not. The example in (31) can clearly show that Oromo compounds can behave just like any simplex words of the language.

(31)	Example	gloss
(a)	Tolaan <i>mana ijaare</i> Tolaa-n <i>mana ijaar-e</i> Tola-NOM house construct-PV Tola house constructed	'Tola constructed a house'
(b)	Tolaan <i>mana kitaabaa ijaare</i> Tolaa-n <i>mana kitaabaa ijaar-e</i> Tola-NOM house book of construct-PV Tola library constructed	'Tola constructed a library'

Looking into (31), one can note that the simplex word written in bold in (31a) and compound in (31b) behave in a similar manner in syntactic construction and one can substitute the other. This can clearly imply that compounds of Oromo can play the role of any simplex words of the language. Moreover, the compound *mana kitaabaa* (*mana* 'house' + *kitaabaa* 'book of') 'library' and the simplex Oromo word *mana* 'house' are also classified in the same syntactic category (noun).

4.1.4 Acronymization and Abbreviation

Abbreviation is a form of lexical elaboration in which new words are coined via the loss of materials from a word; but it differs, however, from truncation and blending in that prosodic categories do not play a prominent role (Mirabela and Ariana 2009:557). The authors also state that abbreviations and acronyms are similar in nature to blends, because both blends and abbreviations are amalgamations of different parts of words. Acronymization is a popular way of forming words in English. An acronymization that is formed via taking the initial letters from the individual words is said to be initialism. Abbreviation is a process of forming a new a word via taking only some parts of existing words. Hence, abbreviation is a short form of a word whereas acronymization is process where new words coined through taking the initial components of words.

Abbreviation and acronymization are used for a variety of reasons such as to avoid repetition, save space, or conform to conventional usage. Various leading authorities, institutions and individuals may not agree on the capitalization or punctuation for many abbreviations and acronyms and therefore, there are no set rules both at a country and at the world level (Mirabela & Ariana 2009:557). This includes various authoritative bodies of Oromia in particular and our country, Ethiopia, in general. The author recommends not to abbreviate unless there is a good reason to do so and the writer knows that the reader will understand. Acronyms are most commonly formed by taking the initial letters of multi-word sequences to make up a new word (Mirabela & Ariana 2009:557). Apart from the words composed of the initial letters; there are also acronyms that incorporate non-initial letters.

Moreover, the pronunciation and spelling of acronyms is also an interesting perspective on the structural properties of abbreviation. For instance, the English acronym FIFO (first in first out), VAT (*value added tax*) and the Oromo abbreviated word *BBO* (*Biroo Barnoota Oromia*) 'Oromia education bureau' are acronymed with capital letters without punctuation. In the case of BBO, it is pronounced by naming individual letter (called initialisms). The first two English words are pronounced by applying regular reading rules just like other words (called acronym). In Oromo abbreviations can be spelled with lowers-case letters, as in, for instance, the Oromo abbreviation *fkn* (*fakkeenya*) 'example'. Here, *fkn* is not pronounced letter by letter, but the unabbreviated form is pronounced.

The following data are collected from Wiirtuu volume 5-11 and Qajeelcha (1998) of the standardization committee of Oromo.

(32) Full form	abbreviation/acronyms	gloss
(a) <i>fakkeenya</i>	<i>fkn</i>	'example'
<i>yokkaan</i> (<i>yokkiin</i>)	<i>ykn</i>	'or'
(b) <i>haloota booda</i>	<i>Dh.b</i>	'AD'
<i>dhaloota dura</i>	<i>Dh.d</i>	'BC'
<i>akka lakkoofsa habashaatii</i>	<i>A.L.H</i>	'Ethiopian calendar'
<i>akka lakkoofsa Awroopaattii</i>	<i>A. L..A</i>	'European calendar'

<i>lakkoofsa saanduqa poostaa</i>	<i>L.S.P</i>	'p.o. box'
<i>waan kana fakkaatan</i>	<i>Wkf</i>	'etc'
<i>waraana adda bilisummaa Oromoo WABO</i>		'army of Oromo liberation front'
<i>dhaabbata dimokratawaa ummatta Oromo DhDUO</i>		'OPDO'

As shown in (32a), the abbreviated forms are composed of one word whereas the data in (32b) are acronyms formed from two or more words. Such kinds of acronyms in Oromo are commonly written both with and without period. In data (32b), except *WABO* <*wabo*> *waraana adda bilisummaa Oromoo* 'army of Oromo liberation front' and *DhDUO* '<ɗadiʔo>' in which they will be read as a word, all the rest forms can be categorized under acronyms that are read letter by letter. On the other hand, acronyms that are formed via taking the initial letters of English words like that of *OPDO* 'Oromo people's democratic organization' and *OBN* 'Oromia broadcast network' are pronounced letter by letter just like any other English abbreviation such as CEO (chief executive officer). In *yokkaan* → *ykn* 'or' and *fakkeenyaaf* → *fkn* 'example' as we can see in (32a), though they are abbreviated, such abbreviations are not read letter by letter; but as a word.

4.1.5 Blending in Oromo

Lieber (2009:52) states that blending is a process of word formation in which parts of lexemes that are not themselves morphemes are combined to form a new lexeme. Blending involves the coinage of a new lexeme by fusing parts of at least two other source words of which either one is shortened in the fusion and/or where there is some form of phonemic or graphematic overlap of the source words, (Gries 2004:639). From the semantics point of view, blending is the word formation process in which parts of two or more words combine to create a new word whose meaning is often a combination of the original words. Hossenzadeh (2014:18) states the term blending as the combination of two forms where at least one has been shortened. The shortening can be by simple omission of a part of a word or it can be a result of overlapping sounds or letter. In the formation of blends, the first element of the first word is added to the second element of the second word. Blending is the formation of new words by combining parts of two words or a word plus a part of another word such as the often quoted blend 'brunch' (breakfast + lunch).

Blending is another productive mechanism of shortening words in Oromo. It is also one of lexical enrichment techniques used by the standardization committee of Oromo in elaborating the language. The formation of blends in Oromo is classified and presented as follows. The following data are collected from the standardization committee of Oromo Wiirtuu volume 5 1992, volume 6 1993, volume 7 1995, volume 8 1998, volume 9 2001, volume 10 2005.

(33)	Combination of words	blending	gloss
	<i>lafa</i> 'earth' + <i>xinxala</i> 'study' →	<i>lafxina</i>	'geography'
	<i>xinxala</i> 'study' + <i>nyaata</i> 'food' →	<i>xinyaa</i>	'dietetics'
	<i>irra</i> 'than' + <i>guddaa</i> 'greater' →	<i>irgu</i>	'greater than'
	<i>dura</i> 'pre' + <i>galaa</i> 'that enter' →	<i>durgaa</i>	'preposition'

As we can see in the above data, lexical enrichment of the language using the blending technique is very productive in Oromo. In (33) existing words are shortened to form anew word. Here, one can note that all the new blends in (33) coined though the reduction from both the first and second constituents. For instance, the blend *lafxina* 'geography' is formed by the property of loss of the second segments of both constituent. In other words, the first segment of each constituent is taken to coin the blend *lafxina* i.e. *laf-* (the first segment of the first word *lafa*) + *xin-* (the first segment of the second word *xinxala*). Such combination can be called reduced form + reduced form which results in a new blend form and can be represented via mathematical expression called $WX + YZ \rightarrow WY$. Here, W and Y are assumed to be the first element of the first word and the first element of the second word respectively.

Oromo blends are also seen to be formed through another type of reduced form + reduced form combination which is different from the formation in (33). The following data are extracted from Qajeelcha (1998) compiled by the standardization committee of Oromo to exemplify mechanism of blending in Oromo stated above.

(34) Combination of words	blending	gloss
<i>agarsiisa</i> 'demonstration'+ <i>hojjaa</i> 'work'→	<i>agoja</i>	'demonstration sight'
<i>barreeffama</i> 'written document' + <i>awwaala</i> 'burial'→ <i>barwaala</i>		'epitaph'
<i>caasaa</i> 'structure' + <i>hima</i> 'sentence' →	<i>caasima</i>	'syntax'
<i>gamisa</i> 'half' + <i>geengoo</i> 'circle' →	<i>gamingoo</i>	'semicircle'
<i>ilaalcha</i> 'view' + <i>addunyaa</i> 'world' →	<i>ilaanya</i>	'ideology'
<i>lama</i> 'two' + <i>hidhii</i> 'lip' →	<i>ladhii</i>	'bilabial'
<i>ogeessa</i> 'expert' + <i>hojii</i> 'work' →	<i>ogojii</i>	'operator'
<i>sagalee</i> 'sound' + <i>jijjiira</i> 'change' →	<i>sajira</i>	'metathesis'
<i>xaxamaa</i> 'complex' + <i>hima</i> 'sentence'→	<i>xaxima</i>	'complex sentence'

As can be seen in (34), a number of Oromo blends can be formed through truncation of both words. Accordingly, the first segment of the first word will be blended with the second segment of the second word. For instance, the blend *barwaala* 'epitaph' is coined via taking the first segment of the first word *barreeffama* 'written document' and attaching it into the second segment of the second word *awwaala* 'burial'. Such type of blend formation seems productively applied in the materials published by the committee. The formation of blends in (34) can, generally, be represented by the expression $WX + YZ \rightarrow WZ$; W and Y as well as X and Z are the first and the second segments of the combining words respectively. The cases of the formations of the blends *agoja* 'demonstration sight' and *ilaanya* 'ideology' are a bit different from the rest of the reduced + reduced formations as the final vowels of the respective blends became short vowels.

Blends in Oromo can also be coined through deleting parts of the first and second constituents as the following data taken from Wirtuu volume 8 and 9 shows.

(35) Noun	noun	blends	gloss
a)	<i>faaya</i> 'cloth' + <i>yeroo</i> 'time' →	<i>faayeroo</i>	'fashion'
b)	<i>gama</i> 'side' + <i>midhaa</i> 'effect'→	<i>gammidhaa</i>	'side effect'
c)	<i>afaan</i> 'language' + <i>dhaba</i> 'who lose'→	<i>afdhaba</i> (<i>af-dhaba</i>)	'language lose'
	<i>afaan</i> 'language' + <i>yaa'ii</i> 'congres' →	<i>afyaa'ii</i> (<i>af-yaa'ii</i>)	'spokesman'

<i>afaan</i> 'language' + <i>maka</i> 'mixed' →	<i>afmaka</i> (<i>af-maka</i>)	'code mixing'
<i>kafaltii</i> 'payment' + <i>mindaa</i> 'salary' →	<i>kamindaa</i>	'payroll'

As it is depicted in (35) another strategy has been applied to coin new blends. In (35b) and (35c), only the first word undergoes truncation and attached to the second word. In (35b), only the last vowel of the first word is omitted and the remaining segment of first word is attached to the second word without truncation as in, for instance, *gama* 'side' + *midhaa* 'effect' → *gammidhaa* 'side effect'. Though the formation of all blends under (35), except (35a), follows reduced word + full word (WX +YZ →WYZ), the formation of blends in (35c) is a bit different as only a few segments (only the first cv or vc) of the first word are taken as compared to blends (35a & 35b). Moreover, as can be seen from the first three blends in (35c), some Oromo blends are seen to be spelled at least in two ways; as a single word and being hyphenated as in, for instance, *afyaa'ii* or *af-yaa'ii* 'spokes man' respectively. The formation of blend *faayeroo* 'fashion' is different in that it followed common consonant structure; *faaya* 'cloth' + *yeroo* 'time' → *faayeroo* 'fashion'.

Generally, internal means of lexical elaboration techniques of Oromo are discussed under this chapter. Compounding, semantic extension, derivation, abbreviation and blending are internal means of lexical elaboration techniques in Oromo. Compounding is extensively used for lexical elaboration of the language. The constituents of compounds in Oromo involve mainly the categories of nouns, adjectives, post-position and verbs. In relation to compound types, nominal, adjectival, adpositional and verbal compounds have been identified in Oromo. Of the three types of compounds, nominal compounds are most productive, followed by adjectival, adpositional and verbal compounds. Both adpositional and verbal compounds are less productive as compared to other methods in Oromo.

When viewed from the semantics perspective, compounds of Oromo can be classified in three; endocentric, exocentric and appositional compounds. Most of the compounds of Oromo we have seen under 4.1.3.1.1 are endocentric compounds where the meaning of the resultant compound emanates from the meanings of its constituents. Exocentric compounds in Oromo are compounds

in which their meanings are not the result of the meanings of their constituents as in, for instance, *garaajab-eessa* ('stomach strong -NMZR) > *gara-jabeessa* 'heartless, hard-hearted', of *tuul-aa* (self pile-NMZR) 'boaster'. Exocentric compounds are not many as compared to endocentric compounds. Appositional compounds are compounds in which both constituents contribute to the meaning of the resultant compound (see section 4.1.3.1.3). Appositional compounds of Oromo are not as productive as endocentric compounds.

Regarding the compound formation in the language, formal changes such as vowel deletion of segments are exhaustively used in Oromo compounds. Following the formal changes like deletion of segments in word boundary and deletion of terminal of the second constituent and the majority of the compounds are written as single lexical items in Oromo. Compounds in the language are written as both separate and single words. However, there are inconsistencies in writing compounds in Oromo. Compounds are seen to be written with hyphens between the constituents, with a space between the constituents and as a single word (see section 4.1.2). These inconsistencies should be harmonized in order to minimize variation in the language. Otherwise, these inconsistencies can highly affect standardization of the language.

As it is discussed under this chapter, semantic extension is also another internal means of lexical elaboration technique applied by the standardization committee of Oromo to enrich the language. As per the data collected from the materials published by the standardization committee of Oromo shows, the meanings are extended from parts of a human body and plants (see section 4.1.1). Abbreviation and blending are also other techniques of internal elaboration of the language. Blending is extensively used in the elaboration of the language. Blending in Oromo is formed via the reduction of segments from at least one of the combining words (see section 4.1.5). The standardization committee of Oromo employed more instances of blending compared to other methods such as abbreviations acronymization which are less productive in Oromo. The acronym forms of two or three words are usually written in Oromo with and without period.

4.2 External means of lexical elaboration

Now a day, no country seems to be an island that isolates itself from the rest of the world. People usually move from place to place. In addition, technology helped human kind to feel as if the

whole world lives in a single village. Hence, there is intercultural contact due to communication. As a result, one of the most easily observable results of intercultural contact and communication is a set of loanwords that are imported into the vocabulary of each language involved (Hoffer 2005:53). The contact could be distant such as via instructions, media and due to foreign language occupations that may yield the entrance of lexical items. Therefore, external means of lexical elaboration is a method by which a language covers its lexical gaps that results from cultural contact and communication. The case of Oromo is not an exception. In this regard, apart from the internal means of lexical elaboration, the standardization committee of Oromo applied external means to cover the lexical gaps of the language. Hence, borrowing and loan translation applied by the standardization committee of Oromo are discussed under this subsection of the chapter.

4.2.1 Borrowing

Borrowing refers to bringing a form from the source language into the recipient language (Bahumaid 2015:15). Borrowing may involve importing and retaining the source language form without substituting any of its elements by recipient-language elements. To Haspelmath (2009:36), borrowing is adopting elements from other languages into the recipient language by the native speakers or an imposition of properties of their native language onto a recipient language by the non-native speakers. Borrowing has also been used in a more restricted sense to refer to the incorporation of foreign elements into the speakers' native language (Haspelmath 2009:36, Winford 2005:377). Although the extent of adaptation depends from language to language, borrowed items are adapted to the phonology, writing system and grammar of the language that borrows them. In this regard, in Oromo, almost all words that are borrowed words from other languages are adapted to the phonology, morphology and orthographic structures of the language as we shall see in the sections to come. Borrowing could be complex and the purpose of the present research is not to discuss the entire issues of concepts and its application. Rather, borrowing in Oromo as a means of lexical elaboration and the process of adaptation of loanwords in the structure of the language is presented in the following sections.

Oromo has borrowed a large number of words from various languages. For instance, as to Griefenow-Mewis (1997:1) it is not unexpected that we can find several Swahili loanwords in Oromo because Swahili and Oromo speaking people were neighbors for, at least, several

centuries. The author further states that if we are looking for Swahili-loanwords in Oromo we have, of course, to examine the southern Oromo-dialects first. Hence, such kind of borrowing is due to language contact. Oromo borrowed a number of lexical items from foreign languages due to the influence of religion and modern civilization. For instance, Oromo borrowed different terms from English mainly to express scientific and technological concepts. Loanword integration mechanisms in Oromo are discussed in the following consecutive subsections.

4.2.1.1 Phonological adjustments while borrowing

4.2.1.1.1 Vowel length

As it is mentioned above, loanwords are usually required to conform to the phonological, morphological and syntactic rules/constraints of a recipient language to the extent that they become integrated into native phonology through a gradual and a complicated process (Paradis & LaCharite 2011:763). However, there are conflicting demands on borrowers to keep the borrowed form as much similar as possible to the source form; but at the same time abide by a recipient language grammar (Kenstowicz 1983:316; Abu Guba 2016:33). These conflicting demands that have to be resolved are preserving as much information from the source word as possible and satisfying the constraints that make the lexical item sound like a word of the recipient language (Kenstowicz 1983:316). Since, almost all loanwords that enter into the lexicon of Oromo passes phonological adaptation, vowel length as one of phonological adjustments in Oromo is discussed under this subsection. The following table shows some of the modified direct and indirect borrowings from various languages the data are collected from the standardization committee of Oromo Wiirtuu volume 10 2005 (see also Bender 1972:320-321).

Table 17: modified foreign loanwords in Oromo

Oromo	source language	source word	Gloss
<i>baakteriyaa</i>	English	<i>bacteria</i>	'bacteria'
<i>gommaa</i>	Amharic<Italian	<i>gomma</i>	'rubber'
<i>kaartaa</i>	Amharic<Italian	<i>karta</i>	'map'
<i>kantiibaa</i>	Amharic	<i>kantiba</i>	'mayor'

As can be seen in table (17), Oromo borrowed words from various languages and adapted those words into its phonology. Most of the above loanwords are adapted into the phonology of Oromo via vowel length word medially and word finally as in, for instance, *kantiibaa* 'Mayor' and *gommaa* 'rubber'. As can be seen from table 17, vowel lengthening is attested for low central vowel a both word medially and word-finally. This does mean, however, that other vowels do not undergo during borrowing.

4.2.1.1.2 Vowel replacement

Vowel replacement is one if the phonological adaptation through which a language adapt non existing sounds with its own sounds during lexical borrowing. Based on the analysis of the phonological adaptation during the incorporation of foreign elements into Oromo, all loanwords in Oromo in the present study undergone phonological modifications as seen in the following sections. The following data under table 18 are collected from Qajeelcha (2008) and Wiirtuu volume 7 (1995) compiled by the standardization committee of Oromo to show the loanword and the phonological adaptation process during borrowing.

Table 18: Adaptation of English vowels in Oromo

English	Oromo	vowel replacement	gloss
/æsid/ →	<i>asiidii</i>	/æ/ → /a/	'acid'
/ælkəhɔ:l/ →	<i>alkoolii</i>	/æ/ → /a/, /ə/ → /o/	'alcohol'
/ka:rbən/ →	<i>kaarboonii</i>	/ə/ → /o/	'carbon'
/bʌdʒɪt/ →	<i>bajata</i>	/ʌ/ → /a/, /ɪ/ → /a/	'budget'
/rɪpʌblik/ →	<i>ripaablikii</i>	/ʌ/ → /a/	'republic'
/mɜ:rkjəri/ →	<i>meerkuurii</i>	/ɜ/ → /e/	'mercury'
/bætəri/ →	<i>baatirii</i>	/æ/ → /a/, /ə/ → /i/	'battery'
/fɔ:rəm/ →	<i>fooramii</i>	/ɔ/ → /o/, /ə/ → /a/	'forum'

As seen in table 18, a number of vowel replacements took place while integrating foreign concepts into Oromo. The vowel /ʌ/ is does not exist in the phonology of Oromo and is replaced by the existing vowel /a/. By the same token, the non existing vowels such as, /æ/, /ɜ/ and /ɔ/ in the phonological formation Oromo are replaced by the existing nearest sounds /a/, /e/ and /o/ as

in, for instance, /æsid/→asiidii, /mɜ:rkjəri/→meerkuurii and /fɔ:rəm/→fooramii respectively. The processes of vowel replacement and vowel length are also attested during the integration of loan words into Hadiyya as it is stated in Samuel (2017:203). The above data on table (18) depicts that in addition to vowel replacement, loan words which end in consonant and vowels at their terminals as in /brʌʃ/→burushii. In this case, the vowel *u* is adapted from the spelling of the loan word *brush* as there is no *u* in the pronunciation of the word *brush*.

Apart from the replacement of non existing vowel by the nearest vowel of Oromo, the non existing diphthongs and triphthongs of English loanwords are adapted into the phonology of Oromo. Before directly discussing the adaptation of diphthongs and triphthongs into the phonology of Oromo, an overview about diphthongs and triphthongs in English is vital.

Diphthong is a term used in the phonetic classification of vowel sounds on the basis of their manner of articulation; it refers to a vowel where there is a single (perceptual) noticeable change in quality during a syllable (Crystal 1997:146). It is a glide or a movement from one vowel to another one in which the first one is prominent than the last one. It is the term used to describe a process where a monophthong has become a diphthong (has been diphthongized) (Crystal 1997:146). Few examples of diphthongs are /aʊ/, /ɔɪ/ as in *how* /haʊ/ and *boy* /bɔɪ/ respectively.

There are also sets of English complex vowels called triphthongs. They refer to a type of vowel where there are two noticeable changes in quality during their production (Crystal 1997:146). There are glides not only from one pure vowel to another, but also from the second vowel to the third, so three pure vowels are involved (Crystal 1997:146). This triphthong starts at the open region, goes to the front close region and ends at the middle of the mouth. Such triphthongs are found in the English words such as *buyer* /bʌɪər/, *fire* /fɪər/ and so on. The tongue too has a complex movement during the articulation.

The English diphthongs and triphthongs do not exist in the phonology of Oromo. Hence, there are various phonological adaptations during the integration of foreign words to Oromo. The following data which is collected from Qajeelcha (1998) and from the standardization committee

of Oromo Wiirtuu volume 10 2005 clearly show the existing phonological adaptations during lexical borrowing.

Table 19: Replacement of English diphthongs in Oromo

English	Oromo	diph(triph)thong replacement	gloss
/dʒiɑ:mətri/ →	<i>ji'oomeetirii</i>	ia → iʔo	'geometry'
/əʊvəri/ →	<i>ovaarii</i>	əʊ → o	'ovary'
fɒ:sfeit/ →	<i>foosfeetii</i>	ei → ee	'phosphate'
/iŋkjubeitər/ →	<i>inkubeetarii</i>	ei → ee, ə → a	'incubator'
/maikrəskəʊp/ →	<i>maayikrooskooppii</i>	əʊ → oo	'microscope'
/reɪʃiʊ/ →	<i>reeshoo</i>	eɪ → e, iʊ → oo	'ratio'
/ziroʊ/ →	<i>zeeroo</i>	I → ee, oʊ → oo	'zero'

As it is depicted in table (19), the English diphthongs or triphthong such as *ou*, *ei*, *əʊ*, *ia*, and *iou* are replaced by existing Oromo simple vowels. For instance, the English *ou* and *əʊ* are replaced by long *oo* in Oromo, as in /reɪʃiʊ/ → *reeshoo* /reeʃoo/ 'ratio' and /əʊvəri/ → *ovaarii/ovaarii* 'ovary', respectively. When words end with consonants, they are adjusted with vowels word finally as in /maikrəskəʊp/ → *maayikrooskooppii* 'microscope', /iŋkjubeitər/ → *inkubeetarii* 'incubator'. The English vowel /ə/ is replaced by /e/ though in some instances is replaced by /a/, /i/, or /o/. Vowel harmony may also result in replacement of /ə/ with /o/.

As shown in (36), when Amharic loanwords are incorporated into Oromo, impermissible consonant cluster in both Amharic and Oromo is adjusted through inserting /i/ and /i/ respectively between the consonant of a word. Consider the following data (36) which are collected from Wiirtuu volume 1-11 published by standardization committee of Oromo.

(36) Amharic	Oromo		gloss
<i>dirama</i> (Amh. < Eng.)	<i>diraamaa</i>	<i>i</i> → <i>i</i>	'drama'
<i>firaʃ</i>	<i>firaashii</i>	<i>i</i> → <i>i</i>	'mattress'
<i>gibir</i>	<i>gibira</i>	<i>i</i> → <i>i</i>	'tax'
<i>kirr</i>	<i>kirrii</i>	<i>i</i> → <i>i</i>	'string'

<i>milaŋ</i>	<i>milaacii</i>	<i>i</i> → <i>i</i>	'razor'
<i>simminto</i> (Amh.<Italian.)	<i>simintoo</i>	<i>i</i> → <i>i</i>	'cement'
<i>ŋiro</i>	<i>shiroo</i>	<i>i</i> → <i>i</i>	'Ethiopian cultural food'

As it is depicted in (36), the Amharic central vowels /i/ and /ə/ are replaced by /i/ and /a/, respectively. But, there is an exception to this general rule. In few loanwords, the vowel /i/ is realized as /u/. The next examples illustrate this fact.

(37) Amharic	Oromo	vowel replacement	gloss
<i>bək'k'ollo</i>	<i>boqqolo</i>	<i>ə</i> → <i>o</i>	'maize'
<i>sik^war</i>	<i>sukkaara</i>	<i>i</i> → <i>u</i>	'sugar'
<i>ŋigut'</i>	<i>shugguxii</i>	<i>i</i> → <i>u</i>	'pistol'

In (37), the vowel /i/ is changed into /u/ due to vowel harmony (see section 3.2). In other words, the vowel /i/ is progressively assimilated to /u/ when it immediately precedes /u/. The central vowel /ə/ is also replaced by /o/ in Oromo through progressive assimilation as in Loanword from Amharic *bək'k'ollo* → *boqqolo* 'maize' *ə*→*o*.

On the other hand, loanwords ending in consonant sounds in the source language are adjusted in Oromo by adding terminal vowels as shown in data (38). The data are collected from the standardization committee of Oromo Wiirtuu volume 1-10 (1985-2010).

(38) Oromo		gloss
(a) <i>arvii</i>	<	'nerve'
<i>eleektiroonikii</i>	<	'electronic'
<i>faayilii/faayila</i>	<	'file'
<i>fiilmii</i>	<	'film'
<i>intarneetii</i>	<	'internet'
<i>neetworkii</i>	<	'network'
<i>oksijiinii</i>	<	'oxygen'
<i>poolisii</i>	<	'police'

<i>seelii</i>	<	'cell'
(b) <i>iskiriinii</i>	<	'screen'
<i>istaadiyamii</i>	<	'stadium'
<i>ispoortii</i>	<	'sport'
(c) <i>daariktara</i>	<	'director'
<i>doolaara</i>	<	'dollar'
<i>kontiraata</i>	<	'contract'
<i>miliyoona</i>	<	'million'

As it is depicted in the above data (38), the English words end in consonants. Therefore, in order to fit to the Oromo word structure, the loanwords receive the terminal vowel /i/ and /a/ (see section 3.2). In the data (38a) the terminal vowel *i*, and in the data (38c), the terminal vowel /a/ are added at the end of the loanwords to fit to the Oromo word structure. On the other hand, in data (38b), /i/ is also added as an epenthetic vowel adjust the impermissible consonant cluster of the language as in 'screen'→*iskiriinii*. Similarly, loanwords from Amharic behave in the same manner. In other words, when words end in consonant in Amharic, they are also adjusted by adding terminal vowel (most commonly, /a/) in Oromo as in sik^war→*sukkaara* 'sugar', gibir →*gibira* 'tax'.

4.2.1.1.3 Consonant adaptation

There are phonemes which were not part of Oromo phonemic inventory. For instance, the phonemes /s'/, /v/, /p/, /z/ and /ʒ/ were not part of Oromo phonemic inventory. Except the last, /ʒ/, all are included in the inventory as loan phonemes by the standardization committee of Oromo (see section 3.1). These phonemes were introduced into Oromo through lexical borrowing to designate mainly scientific and technological concepts. Therefore, the non-existent consonants such as /s'/ <ts>, /v/ <v>, /p/ <p>, /z/ <z> and /ʒ/ <zh>, <zn> or <zy> were written without replacement in the target language (Oromo), as shown in (39). The following data were taken from the standardization committee of Oromo Wiirtuu volume 1-11(1985-2010 E.C).

(39) Oromo	source language	gloss
<i>ezyiiyaa/ Ezhiiyaa</i>	English	'Asia'
<i>fiiziksii</i>	English	'physics'
<i>gaazeexaa</i>	Amharic	'newspaper'
<i>Indooneezyiiyaa/Indooneezhiiyaa</i>	English	'Indonesia'
<i>muuza/muuzii</i>	Amharic	'banana'
<i>narvii</i>	English	'nerve'
<i>pirootinii</i>	English	'protein'
<i>poolisii</i>	English	'police'
<i>poozativii</i>	English	'positive'
<i>televiizhina /televiizyinii/</i>	English	'television'
<i>tsabala</i>	Amharic	'holly water'
<i>vaayirasii</i>	English	'virus'
<i>varteksii</i>	English	'vertex'
<i>viizyuuwaala/ viizhuuwaal</i>	English	'visual'
<i>Vitaaminii</i>	English	'vitamin'

As we can see in the above data (39), the occurrence of /p/, /v/ and /z/ is very frequent in loanwords but /s'/ and /ʒ/ is very limited. In addition, loanwords which include the consonant /ʒ/ and /s'/ could occur in free variation with /dʒ/, and /t'/ as in *teelevijiinii /teelevidʒiinii/* 'television' or as in *xabala /t'abala/* respectively. Regarding, phoneme /ʒ/, which is not part of Oromo phonemic inventory currently is being represented by various graphemes like <zh>, <zy>, <zn> and sometimes replaced by the nearest phoneme <j>/dʒ/. The standardization committee of Oromo still has not yet considered the case seriously. It is not either represented with appropriate grapheme. In the written form of the language the non-native consonants seem to be the preferred one.

Segment or syllable deletion is also attested in Oromo during borrowing. Consonants such as /h/ and /j/ tend to get lost between vowels when Amharic loanwords are incorporated into Oromo. For instance, in the borrowing processes of Amharic words into Oromo like */bahirzaf/* → *barzaafii* 'eucalyptus tree' (there is a native word called *bargamoo*), */mahitəm/* → *maataba*

‘seal’, /mäs’ihaf/ → *macaafa* ‘related to holly book’ etc. deletion of /h/ or /hi/ has occurred in order to adjust to Oromo phonological systems. This case is also attested in Hadiyya as it is stated in Samuel (2017:208). The word *macaafa*, which is used in the Oromo Bible translation, the phoneme /s’/ is replaced by <c> /tʃ/. As opposed to this issue, the phoneme /s’/ <ts> has been acquired and accepted as the loan phoneme of Oromo by the standardization committee of Oromo. This loan phoneme is represented by the grapheme <ts> (see section 3.1).

Moreover, there are also consonant cluster adaptations during the inclusion of foreign concepts into the language. Oromo disallows consonant cluster word initially. As a result of this fact, Oromo makes phonological adjustments during the integration of foreign lexical items into its lexicon. Accordingly, loanwords of English which begin with consonant clusters of two or more, word initially are adapted in Oromo phonological systems by inserting a syllable that contain a glottal stop and /i/ word initially (see section 3.2). Griefenow-Mewis (2001:1) states that loanwords have to be adapted to the syllable structure of Oromo. A sequence of two consonants at the beginning of a word has to be separated by either a prothetic or an epenthetic vowel which breaks up the cluster into two syllables. The author gives practical examples from Oromo loanwords. The data in (40) show the adjustment of impermissible word initial consonant cluster in Oromo. The following words are collected from the standardization of Oromo Wiirtuu volume 1-11 (1985-2010 E.C).

(40)	Oromo	gloss
	isfeerii /ʔisfeerii/	'sphere' /sfɪr/
	iskaarvii/ʔikaarviii/	'scurvy' /skɜ:rvi/
	istoomaataa/ʔistoomaataa/	'stomata' /stəʊmətə/
	islayidii /ʔislajidii/	'slide' /slaid/
	ispoortii /ʔispoortii/	'sport' /spɔ:rt/

Looking into the data in (40), one can note that based on the phonotactic constraints of Oromo, word initial consonant clusters in English loanwords such as *sf-*, *sk-*, *sl-*, *sp-*, and *st-* are broken via the insertion of an epenthetic vowel (*i*) to conform to the phonology of Oromo. This is due to the fact that Oromo disallow consonant cluster word initially and word finally (see section 3.2).

Such phonotactic constraints also exist in Hadiyya and are adjusted with the same mechanism as it is stated in Samuel (2017:209).

There are also other mechanisms by which impermissible consonant cluster of English loanwords can be adjusted other than the insertion of an epenthetic vowel (*i*) word initially. The following data collected from Qajeelcha (1998) and standardization committee of Oromo Wiirtuu volume 10 (2005) for an illustration.

41) Oromo	English
<i>pilaastara</i>	'plaster' /plæstər/
<i>pilaatiniyamii</i>	'platinum' /plætɪnəm/
<i>pireezidaantii</i>	'president' /prezɪdənt/
<i>piroojektii</i>	'project' /prɑːdʒekt/

As can be seen from (41), word initial consonant cluster of loanwords can also be conformed to the phonology of Oromo via inserting the vowel *i* in between the consonants *p* and *r* as well as *p* and *l* as in, for instance, *pireezidaantii* 'president' /prezɪdənt/ and *pilaastara* 'plaster' /plæstər/ respectively.

4.2.1.4. Superfluous borrowing

Lexical borrowing could be due to either need or prestige. Internal need for borrowing may result from an emerging socio-cultural change. Borrowing of lexical items could also be due to the existence of a prestige and powerful (from cultural points of view) language that become sources for loans in other languages (Haspelmath 2009:36). In some other cases a language may integrate loanwords from other languages; even though there are its own native words to express a concept. In this case, foreign lexical items are integrated without any need. The following example shows some of foreign lexical items which are unnecessary borrowed though there are genuine Oromo words in the language. The data are taken from Qajeelcha (1998) and standardization committee of Oromo Wiirtuu volume 5 (1992), volume 7 (1995), volume 8(1999), volume 10 (2005).

(42) Superfluous borrowing from English/Amharic

Loanword	source language	genuine word	gloss
<i>cinqamuu</i>	Amharic	<i>hiraaruu</i>	'misery'
<i>egzibishiinii</i>	English	<i>daawii/ agarsiisa</i>	'exhibition'
<i>haakara</i>	English	<i>uggurtaa</i>	'hacker'
<i>kasaaraa</i>	Amharic	<i>honga'a</i>	'bankrupt'
<i>laaptoopii</i>	English	<i>koyaa</i>	'laptop'
<i>maarkii</i>	English	<i>qabxii</i>	'mark'
<i>riipoortara</i>	English	<i>gabaasaa</i>	'reporter'
<i>sheerii</i>	English	<i>qoodachuu/yagutooma</i>	'share'
<i>taabii</i>	English	<i>caancala</i>	'tab'
<i>wakiila</i>	Amharic	<i>bakka bu'aa</i>	'representative'

As the above data (42) depict, there are unnecessary borrowings of words; eventhough there are genuine words in Oromo that can express the given concept. This may affect standardization of the language.

4.2.1.5 Source languages for lexical borrowing in Oromo

A language may borrow lexical items from other indigenous and exogenous languages to express concepts which are beyond its culture.

4.2.1.5.1 Direct borrowing in Oromo

Apart from the indirect borrowing of words from other languages, Oromo has borrowed words directly from other languages. The following data are collected from Wiirtuu volume 5 Qajeelcha 1998 to show the existence of directly borrowed words in Oromo.

43) Amharic	Oromo	gloss
<i>ባቄላ/bakela</i> →	<i>baaqeela</i>	'bean'
<i>ጫማ /f'amma/</i> →	<i>cammaa</i>	'shoe'
<i>ፊሻካ /fj/ka/</i> →	<i>fishkaa</i>	'whistle'

እዳ /ida/ →	<i>idea</i>	'debt'
ክትፎ /kitfo/→	<i>kitfoo</i>	'minced meat'
መከረ /məkər/→	<i>makara</i>	'harvest'
መስቀል /məskəl/→	<i>masqala</i>	'cross'
ሽቶ /šitto/→	<i>shittoo</i>	'perfume'

Languages may borrow in order to enrich its lexicon. Oromo borrowed loanwords to express various concepts from various languages. But, in some of the cases, it is not easy to recognize the exact source languages, particularly, when particular loanwords exist in various languages. The following sections discuss the direct source of borrowing in Oromo.

4.2.1.5.1.1 Loans from Swahili

As to Griefenow-Mewis (1997:157), there are several Swahili loanword in Oromo as a result of the fact that Swahili and Oromo speaking people were neighboring people, for at least several centuries in the southern part of Ethiopia. The following data are collected from Qajeelcha (1998) and Griefenow-Mewis (1997:157) to show some of Swahili loanwords in Oromo.

(44) loanword (Swahili)	Oromo	gloss
<i>elfu</i>	<i>elfuu</i>	'thousand'
<i>gari</i>	<i>gaari</i>	'cart'
<i>jembe</i>	<i>jembee</i>	'hoe'
<i>kiti</i>	<i>kiiti</i>	'chair'
<i>mnada</i>	<i>munada</i>	'auction'
<i>msinga</i>	<i>misinga</i>	'millet'
<i>shoka</i>	<i>sookaa</i>	'axe'
<i>simu</i>	<i>simu</i>	'telephone'

Due to language contact between Swahili and Oromo, loanwords enter into Oromo. Though there are some Swahili loanwords like *elfu* 'thousand' that found in other language (in Amharic as in ሰልፍ /ilf/ too, there are also some Swahili loanwords like *munada* 'auction', *jembee* 'hoe' which are found only in Oromo. This may indicate that Swahili loanwords have been incorporated into

Amharic via Oromo as there was contact between Swahili and Oromo speaking people, particularly in the southern and south-eastern Ethiopia (Griefenow-Mewis 1997:157).

4.2.1.5.1.2 Loans from Arabic

Oromo borrowed a number of words from Arabic to elaborate its lexicons. As a result of cultural contact (through the Islam or trade) and a number of Arabic words entered into the Oromo lexicon (Griefenow-Mewis 1997:157). Leslau (1959: 275), also states that Arabic loanwords come in the Easter Oromo due to the religious background (who are Moslems) and because of their commercial contact with Moslems and Arabs. There are cases in which Arabic loanwords found only in Oromo and whose roots are not found in Amharic. This may be one of the reasons that make us to think that some Arabic loanwords might have entered into Oromo without an intermediary language. The following are some of the examples of direct loans from Arabic. The data are taken from Standardization committee of Oromo Wiirtuu volume 4 (1991) and Leslau (1959:287) (see also Samule 2017:214).

(45) Arabic	Oromo	gloss
<i>dʒimat</i> →	<i>Jimaata</i>	‘Friday’
<i>dʒinn</i> →	<i>jinnii</i>	‘spirit’
<i>ḥams</i> →	<i>Kamisa</i>	‘Thursday’
<i>kataba</i> →	<i>katab-</i>	‘write’
<i>ʕumr</i> →	<i>umurii</i>	‘age’

As shown in (45), the days of a week like *dʒimat*→*Jimaata* ‘Friday’, *ḥams*→*Kamisa* ‘Thursday’ show that Oromo included Arabic loanwords in its lexicon via direct borrowing as there was cultural contact of Arabs with easter Oromo due trade and religion (Griefenow-Mewis (1997:158). The Arabic loanword such as *kitaab-* ‘write’ is productively employed in the designation of concepts related to print through derivation in Oromo, as in *katabaa*‘writer’, *kitaaba* ‘book’. For days of the week see Ali & Zaborsk (1990:23) in which some of the names of the week are borrowed from Arabic.

Though the task of identifying the exact direction of borrowing may be very complicated, it is obvious that there are several Arabic loanwords that occur in both Oromo and Amharic. However, it could be assumed that such Arabic loanwords were incorporated into Oromo either directly (as Oromo speakers have a contact with Arabic speaking peoples) and indirectly via Amharic. As per Griefenow-Mewis (1997:158), there were different ways for originally Arabic words to enter in the Oromo lexicon, sometimes directly from Arabic (through the Islam or trading), sometimes via Amharic and, surely, sometimes through Swahili.

Arabic loanwords that exist in both Oromo and Amharic are depicted as follows. The data are taken from Qajeelcha (1998) Wiirtuu volume 7 and Leslau (1959:285). These loanwords are also stated to have been borrowed by Amharic and Hadiyya (Samuel 2017:216).

(46) Oromo	Amharic	gloss
<i>adaba</i>	<i>ʔadəb</i>	‘good manners’
<i>akiima</i>	<i>hakim</i>	‘physician’
<i>alkoolii</i>	<i>ʔalkol</i>	‘alcohol’
<i>farasullaa</i>	<i>fərəsulla</i>	‘items of 17 kg’
<i>irsaasii</i>	<i>(i)rsaas</i>	‘lead’
<i>jinjibila</i>	<i>zindzibil</i>	‘ginger’
<i>jabanaa</i>	<i>dʒəbəna</i>	‘jug or kettle of coffee pot’
<i>kiisii</i>	<i>kis</i>	‘pocket’
<i>kooraa</i>	<i>koriffa</i>	‘saddle’
<i>maqasii</i>	<i>məqəs</i>	‘scissors’
<i>masgiida</i>	<i>məsgid</i>	‘mosque’
<i>miizana</i>	<i>mizan</i>	‘balance’
<i>qishirii</i>	<i>qifir</i>	‘coffee grain without peel’
<i>qurunfudii</i>	<i>qirunfud</i>	‘clove’
<i>saamunaa</i>	<i>samuna</i>	‘soap’
<i>sa’atii</i>	<i>səʔat</i>	‘watch’
<i>waraqata</i>	<i>wəraqət</i>	‘paper’

As it is shown in the above data, there are a number of Arabic loanwords in Oromo that Oromo used for its lexical elaboration. But, what is difficult to clearly states is the direction of borrowing. Some of these and other loanwords of Arabic in Oromo might have been borrowed directly from Arabic itself as there are contacts of these peoples in eastern, south east, western part of the country (Leslau1959:285; Griefenow-Mewis 1997:8). On the other hand, some of the loanwords of Arabic in Oromo might have entered into Oromo indirectly via Amharic as Amharic had got higher status (as a linguafranca, medium of instruction, and language of administration).

4.2.1.5.2 Indirect loans in Oromo

Perhaps the most frequently encountered products of cultural contact are the set of loanwords that follow from intercultural communication. English and Amharic are the most important sources of borrowing for Oromo lexical elaboration. In fact, the majority of English loanwords entered Oromo via intermediary language, Amharic. The main reason for this is that Amharic had an opportunity to be developed for a longer period of time as compared to Oromo. In other words, Amharic enriched its lexicon for different scientific and technological concepts by incorporating mainly English loanwords. The development of Amharic as a language of administration, media, and medium of instruction provides a model for the development of other Ethiopian languages (Samuel 2017:214). Therefore, it is possible to say that the majority of foreign loanwords (e.g., English, Italian and French) have been incorporated into Oromo, mainly, via Amharic. But, it cannot be concluded that all foreign loanwords are incorporated via Amharic as there are direct loans from other foreign languages like that of Swahili and Arabic as a result of language contact in southern and other sections of the country (Griefenow-Mewis 1997:8). In addition to the role of Amharic as an intermediary language, it is also possible to perceive that some of the English loanwords are introduced into Oromo through the educated bilinguals, particularly in the school textbook development and through various media in general.

4.2.1.5.2.1 Indirect loan from Italian

Apart from the direct borrowing of words from other languages, Oromo integrated loanwords of other languages in its lexicon via indirect borrowing. Amharic is an intermediary language that became a source language for the enrichment of Oromo lexicon. Several Italian words in Oromo

via Amharic as Amharic was a language of education, administration, medium of instruction and was the only federal working language for a long period of time. Only one instance is attested as direct a loan, i.e., *abokaatoo* ‘lawyer’ which is not attested in Amharic as a loanword. This Italian loanword is also found in Hadiyya as stated in Samuel 2017:219). In other cases, as can be seen from (47), the Italian loanwords were borrowed into Oromo via Amharic. The data are taken from Qajeelcha (2008) (see Samuel 2017:219; Bender 1972:320-321).

47)Source language	Amharic	Oromo	Gloss
<i>avocato</i> →	---	<i>abokaatoo</i>	‘lawyer’
<i>borsa</i> →	<i>borsa</i> →	<i>boorsaa</i>	‘bag’
<i>cemento</i> →	<i>siminto</i> →	<i>simintoo</i>	‘cement’
<i>dzesso</i> →	<i>dzesso</i> →	<i>jeessoo</i>	‘gypsum’
<i>pasta</i> →	<i>pasta</i> →	<i>paastaa</i>	‘pasta/spaghetti’
<i>piazza</i> →	<i>piyassa</i> →	<i>piyaassaa</i>	‘piazza’ (place name)
<i>pizza</i> →	<i>piza</i> →	<i>pizaa</i>	‘pizza’
<i>posta</i> →	<i>posta</i> →	<i>poostaa</i>	‘post’
<i>spago</i> →	<i>sibago</i> →	<i>sibaagoo</i>	‘string’
<i>veranda</i> →	<i>bərənda</i> →	<i>barandaa</i>	‘veranda’

All of the above Italian loanwords are included in the list of lexical items of Oromo and are in circulation. These lexical items are already incorporated within the linguistic system of Oromo and are proven to be found in the materials compiled by the standardization committee of Oromo and in the various primary and secondary textbooks of the language. This is unquestionable. Rather, what is more challenging is to clearly and fairly state whether Italian loanwords are direct loans or indirect loans. We can, however, conclude, based on the data in (47) that these Italian loanword to have been entered into Oromo via Amharic for perhaps two reasons. First as it has been mentioned above, Amharic was a lingua franca, a language of education, administration, medium of instruction and was the only federal working language for a long period of time. Second, since almost all Italian loanwords that are attested in Oromo are also found in Amharic, it is fair to conclude that the borrowing was indirect. Just like any other

loanwords from else languages, Italian loanwords in Oromo conform to the phonological constraints of the language as in, for instance, *bərənda* > *barandaa* ‘veranda’ (see 4.2.1.1 & 4.2.1.2).

4.2.1.5.2.2 Indirect loan from Greek

Oromo had also elaborated its lexicon through borrowing mainly religious words from Greek via Amharic. Some of loanwords from Greek are depicted under the following illustration.

48) Greek	Amharic	Oromo	gloss
<i>diákonos</i> →	<i>dīyaqon</i> →	<i>daaqonii</i>	‘deacon’
<i>christianós</i> →	<i>kristiyan</i> →	<i>kiristaana</i>	‘christian’
<i>sapoúni</i> →	<i>samuna</i> →	<i>saamunaa</i>	‘soap’
<i>trapézi</i> →	<i>t’ərəpp’eza</i> →	<i>xarapheezaa</i>	‘table’
<i>tóxo apó tin</i> →	<i>tabot</i> →	<i>taaboota</i>	‘arc of covenan’

4.2.1.5.2.3 Indirect loan from French

French is also another source of borrowing words for Oromo lexical elaboration in particular and the enrichment of the language in general. The following illustrations in (49) are the loanwords from French which entered indirectly into Oromo via Amharic (see also Bender 1972:320-321; Samuel 2017:192)

49)Source language	Amharic	Oromo	Gloss
<i>champion</i> →	<i>fampiwon</i> →	<i>shaampiyoona</i>	‘champion’
<i>chemise</i> →	<i>fəmi</i> →	<i>shamiizii</i>	‘shirt’
<i>chauffeur</i> →	<i>fufər</i> →	<i>shufeerii</i>	‘driver’
<i>vapeur</i> →	<i>babur</i> →	<i>baabura</i>	‘train’
<i>bureau</i> →	<i>biro</i> →	<i>biiroo</i>	‘office’
<i>bicyclette</i> →	<i>biskilet</i> →	<i>biskileetii</i>	‘cycle, bicycle’
<i>cassette</i> →	<i>kasset</i> →	<i>kaaseetii</i>	‘cassette’
<i>cravate</i> →	<i>kiravat</i> →	<i>karabaataa</i>	‘tie’

<i>colonel</i> →	kolenel→	<i>koleneel</i>	‘colonel’
<i>coupon</i> →	<i>kuppon</i> →	<i>cuupoonii</i>	‘coupon’
<i>médaille</i> →	<i>medaliya</i> →	<i>meedaaliyaa</i>	‘medal’
<i>million</i> →	<i>miliyon</i> →	<i>miliyoona</i>	‘million’
<i>parlement</i> →	<i>parlama</i> →	<i>paarlaamaa</i>	‘parliament’

As can be seen in (49), there are several French loanwords that entered into the lexicon of Oromo via an intermediary language (Amharic). This is due to the fact that, as it has been mentioned earlier, Amharic was a language of administration, medium of instruction, media, and an official working language for several years in the country. Oromo applied a phonological adaptation while indirectly borrowing loanwords of French from Amharic such as the replacement of the non existing vowel *i*→*i* and *i*→*a* as in *biskilet*→*biskileetii* ‘cycle/ bicycle’ and *kiravat*→*karabaataa* ‘tie’ respectively.

4.2.2 Loan translation

Loan translation (calque) is the transfer of lexical or grammatical meaning from a model language into a replica language whereby the latter replicates a formal expression typically via translation (Heine & Kuteva 2006:23). In loan translation (calque), composite form is translated directly, element by element. In such ways a new composite form based on indigenous elements is created to translate a foreign concept. Such method of lexical elaboration is productive in Oromo. From the point of view of Sager & Nkwenti-Azeh (1989:23), loan translation can be categorized into literal and conceptual. Literal loan translation refers to the word-for-word replacement of the lexical components of a foreign term into a host language.

4.2.2.1 Literal loan translation

As it has been stated in the previous section, loan translation is one of the productive ways of designation of foreign concepts in Oromo. Loan translation could be literal or non-literal loan translation. With regard to this, a number of technological, mathematics, social science, political, economic, administrative, etc. terms are designated by loan translation method. The following illustration are examples of literal loan translation from English: The following data under (50) and (51) are collected from Wiirtuu volume 10 2005 of the standardization committee of Oromo.

(50) English	literal loan translation	gloss
annual income	<i>galii waggaa</i> (income + of year)	'annual income'
cultural value	<i>sona aadaa</i> (value + of culture)	'cultural value'
free market	<i>gabaa bilisaa</i> (market + of free)	'free market'
inferential meaning	<i>hiika galchaa</i> (meaning + of inferencial)	'inferential meaning'
job description	<i>maalmaallee hojii</i> (description + of job)	'job description'
material culture	<i>wanta adaa</i> (material + culture)	'material culture'
potential energy	<i>human kuufamaa</i> (energy + of stored)	'potential energy'
process approach	<i>adeemsa maleenyaa</i> (process +of approach)	'process approach'
profit tax	<i>gibira bu'aa</i> (tax + of income)	'income tax'
skin cell	<i>seelii gogaa</i> (cell + of skin)	'skin cell'
universal legacy	<i>kennaa hundamtaa</i> (legacy + of universal)	'universal legacy'

As the above data in (50) shows, the standardization committee of Oromo used literal translation method for lexical elaboration of the language. Accordingly, lexicons of indigenous elements are created to translate a foreign concept as in *imaammata afala* (policy + language + outside) 'exoglossic language policy', *human kuufamaa* (energy + potential) 'potential energy', *wanta adaa* 'material + culture') 'material culture'. As can be observed from the collected data on loan translation, in the majority of the cases, literal loan translation is predominantly used followed by non-literal (conceptual) loan translation. The committee also made a literal loan translation from Amharic to elaborate the language. Some of the illustrations are depicted as follows. The data in (51) are collected from Qajeelcha 2008 and from Wiirtuu volume 10 2005.

(51) English	Amharic loan	literal loan translation in Oromo	gloss
auxiliary view	<i>səjaf gəs'ta</i>	<i>mi'uu shaffaxaa</i> (auxiliary + view)	'auxiliary view'
consanguinity	<i>jəsiga zimidinna</i>	<i>firooma dhigaa</i> (relative + of blood)	'consanguinity'
list of items	<i>jəɲik'a ɲajinət</i>	<i>akaakuu meeshaa</i> (type + of item)	'type of item'
pole vault	<i>mirikuz zılaju</i>	<i>taalcha hirkoo</i> (jump + of pole)	'pole vault'
prison	<i>ɲisr bet</i>	<i>mana hidhaa</i> (house + of prison)	'prison'
rightist	<i>k'əɲɲ kinif</i>	<i>mirgalessa</i> (rightist)	'rightist'
shipping line	<i>jəbahir tıranzit</i>	<i>cahii galaanaa</i> (transit + of ocean)	'shipping line'

The above data shows that, Amharic is also used as a source for literal loan translation to elaborate Oromo vocabulary in addition to English. In Oromo the word *mana* 'house' is very productive to designate various concepts. This word can be used productively to elaborate the language internally (compounding) and externally (loan translation). The following data which are collected from Wiirtuu volume 7-11 of the standardization committee of Oromo clearly shows this fact (see appendix H).

(52) Amharic	Oromo	gloss
<i>foto bet</i> (photo + house)	<i>mansuuraa</i> (house + of photo)	'photograph house'
<i>hakim bet</i> (physician+house)	<i>manayaalaa</i> (house+of physician)	'hospital'
<i>kiraj bet</i> (of rent +house)	<i>manakiraa</i> (house + of rent)	'rent house'
<i>migiḅ bet</i> (food+ house)	<i>mansooraa</i> (house +of food)	'restaurant'
<i>məgazən</i> (ware house)	<i>mankuusa</i> (house + of storing)	'ware house'
<i>mikir bet</i> (counsel + house)	<i>mana maree</i> (house + counsel)	'parliament'
<i>məzigəb bet</i> (record + house)	<i>managalmee</i> (house +of record)	'record office'
<i>mədihaniḅ bet</i> (medicine+house)	<i>manaqoriichaa</i> (house + of medicine)	'pharmacy'
<i>posita bet</i> (posta +house)	<i>manapoostaa</i> (house + of posta)	'post office'
<i>s'əgur bet</i> (hair + house)	<i>manarifeensaa</i> (house + of hair)	'barber shop'
<i>səriḅ bet</i> (of wedding + house)	<i>manacidhaa</i> (house +of wedding)	'house of wedding'

As the above data shows, a number of new concepts can be designated using the productive first constituent word *mana* 'house' and other second constituent. Eventhough, most of the terms formed using the first constituent word *mana* 'house' are structurally compound words, they are basically translated from Amharic. This case is also true for Hadiyya in which the second constituent of the compound word is *mine* 'house' combine with other first constituent of compound word form translated compound words as stated in Samuel (2017:232).

4.2.2.2 Non literal loan translation

Other than the word-for-word translation, a number of concepts of the English terms have been translated into Oromo to elaborate the language. The concepts which are nonliterally translated

from English into Oromo are formed through putting lexical items together. This fact is depicted in the following data.

(53) English	non-literal loan translation	gloss
characterization	<i>ijaarsa fakkaattii</i> (construction + character)	'characterization'
consonant cluster	<i>irrabuta</i> (irra 'on' + buta 'of pulling')	'consonant cluster'
logistics	<i>galaafi hidhanno</i> (held for later and + equipment')	'equipment held for later'
suspence	<i>qalbii raasaa</i>	'suspense'

The above data in (53), non literal loan translation is one of the language elaboration techniques that are being used in Oromo. Most of the above single English words of have been translated none literally using two or more Oromo words.

4.3 Summary

In examining the lexical elaboration techniques used by the standardization committee of Oromo I came across two types of lexical elaboration applied by the committee; internal and external means. Of the internal means, meaning extension, derivation, compounding, abbreviation and blending are applied by the committee to elaborate the language. Derivation, compounding and blending are among the extensively applied means through which the standardization committee of Oromo enriched the lexicon of the language. The committee also used borrowings as method of lexical elaboration in addition to exploiting language internal resources to expand its vocabulary as exhausting language internal resources should be given priority before using other means of language enrichment. As the loanwords collected from the materials published by the committee depict, Oromo has borrowed words from various languages through both direct and indirect methods and integrated them into its phonological systems.

The loanwords are adapted into the language through modification. When loanwords are integrated into Oromo, non-existent vowels are subjected to replacement with the nearest vowels. Likewise, diphthongs which are existent in English loanwords will be replaced either by simple vowels or by a glide insertion in the language. Due to the massive borrowing for the designation of scientific, technological, religious and other concepts, some of the non-native consonants are

introduced to Oromo. For example, the phonemes /p/, /v/, /ʒ/, /s/ and /z/ were not part of the phonological system of Oromo. But they are incorporated via loanwords. Furthermore, the phoneme /ʒ/ has not yet been represented by any grapheme officially and it is being represented by various graphemes like using <zy>, <zh>, <J > and /zn/ to write foreign words. Such variation may affect standardization of the language.

On the other hand, the word initial consonant cluster is not permissible in Oromo. Hence, such impermissible word initial consonant clusters are adjusted by epenthetic vowel *i* in loanwords. Superfluous borrowings are attested despite the existence of the native equivalents in Oromo.

Regarding source language preferences, Oromo tends to borrow significantly from English followed by French, Arabic, Swahili and Italian while borrowing from foreign language sources (see section 5.2.1.3.1.3 and appendix E). Loanwords from Greek are not many as compared to other source languages. Amharic and English are the main sources of lexical borrowing for Oromo. This may be due to the fact that Amharic has been serving as an intermediary language for Oromo's borrowing from foreign languages such as English, Italian, French and Greek. But, Swahili loanwords are incorporated in Oromo through direct borrowing as there was contact between Swahili and Oromo speaking people; particularly in the southern and south-eastern part of Ethiopia (Griefenow-Mewis 1997:157).

Regarding loan translation in Oromo, it is one of the methods being extensively used to elaborate the language. Literal loan translation is primarily used in Oromo followed by non-literal or conceptual loan translation. Non-literal loan translation can create better understanding of designated expressions for native speakers as concepts are translated into a target language. In some of the cases, there exists an overlap between the method of loan translation and compounding in Oromo. In Oromo, the word *mana* 'house' is used productively to coin several concepts through loan translation. This productive pattern functions as a loan translation for Oromo. Even though the translated form has similar in structure with Oromo compound words, semantically it is identified as loan translation.

All in all, the standardization committee of Oromo has codified and normalized some of the language standardization related issues discussed under the third and the fourth chapters. The committee also positively contributed to the standardization process of the language via enriching its lexicon as can be seen from the fourth chapter. However, although there are a number of issues that have to be codified and normalized (including alphabet order of the Roman based script, word spacing, short forms of ordinals, how compounds should be written etc.) the standardization committee of Oromo almost limited to codifying words. Even, the words that are codified are not many as compared to what are remaining. Scientific, technological and technical terms are not being given due attention by the committee. Rather, individual writers such as textbook compilers are coining words to designate scientific and technical terms in their own ways. Technical terms are not coined, normalized and distributed from one center. Hence, a separate group of professionals within the standardization committee of Oromo has to work on the codification and normalization processes of scientific and technical terms as such area is almost untouched by the authority. There are also problems from the standardization committee of Oromo in working hand in hand with stakeholders of standardization of the language. Some of what have been codified is not seen to be implemented by, for instance, education sector (textbook writers). There is almost no communication between the authority and the education sector (See appendix G, FGD 3, 28). Generally, the language is in the process of being standardized. The process is, however, very steady.

CHAPTER FIVE

THE ROLE OF EDUCATION TO THE (DE)STANDARDIZATION OF OROMO

This section of the dissertation discusses the role that the education sector played in the standardization process of the language. Hence, the chapter begins with the role education sector (particularly the textbook compilers) played for and against the standardization of the language and end up with the comparison between lexical elaboration by the committee and textbook writers.

5.1. Introduction

As the language becomes used in a greater and greater variety of functions, it becomes more and more important that a near-uniform variety should be available to fulfill all these functions. Just as the proliferation of varying coinages or weights and measures is dysfunctional, so a proliferation of different forms of the language would be highly undesirable in a society that requires widespread communications (Milroy and Milroy 1985). This is also true for the education sector. Variation is not expected in education material such as in textbooks of primary and secondary schools.

5.2 The role of education in the (de)standardization of Oromo

The language policy of the federal democratic republic of Ethiopia states that government institutions, educational institutions and the mass media shall have special responsibility to maintain the standards of Ethiopia's writing systems that are currently in use (MoCT 2020:23). The policy gives due attention to the standard issues of languages of the country. This can include the implementation of regularized (normalized) in the education sector, particularly in the primary and secondary school textbooks. As the standard variety is the choice for education, its use has become naturally associated with being educated and/or intelligent. Hence, the role of textbook compilers in the standardization process of Oromo is described in the sections to come.

5.2.1 Lexical elaboration techniques

Language can be enriched through various strategies; using both the internal and external means. Such lexical elaborations of Oromo in the primary and secondary textbooks are presented in the sections to come.

5.2.1.1 Internal means of lexical elaboration in the textbooks

This section of the chapter discusses the internal sources of lexical enrichment strategy of Oromo in the primary and secondary textbooks. Semantic extension, redefinition, derivation, compounding, blending, abbreviation are ways of lexical elaborations that are employed to enrich the vocabulary of Oromo. Such methods are discussed in detail in the following sections.

5.2.1.1.1 Meaning extension

Meaning extension is an important mechanism by virtue of which new meanings are produced. Guided by the principle of economy, language users reanalyze the meanings of original expressions, their semantic focus and the construction of their components, thus get a different explanation of the original word based on a new understanding of their literal meanings, which eventually lead to the semantic evolution. In the following subsections, meaning extensions from a human body parts, meaning extension from animals' bodies and farming activities, from parts of plants, food as well as from materials and shelters.

5.2.1.1.1.1 Meaning extension from human body parts

In Oromo, meaning extension from a human body parts has been used as mean of lexical elaboration in the education sector. The method has been used by primary school textbook developers during the preparation of the books so as to enrich the lexical gaps of the language. Illustrative examples of this method of lexical elaboration are shown in (39) below. The following data are collected from Grade 7 and 8 biology, mathematics, civics and Oromo textbooks that are published in 2013.

(54) Oromo	meaning 1	meaning 2
<i>adda</i>	forehead	front/ luck/ (military)/political organization
<i>afaan</i>	mouth	language
<i>arraba</i>	tongue	insult
<i>fuula</i>	face	page
<i>fuunyaan</i>	nose	close to something
<i>gurra</i>	hear	who hears
<i>handhura</i>	navel	center/core
<i>harka</i>	hand	percent
<i>ija</i>	eye	important
<i>kofa</i>	armpit	angle
<i>onnee</i>	heart	mental acuteness
<i>quba</i>	finger	information
<i>sabbata</i>	waist	middle (of earth)
<i>hojjaa</i>	height of someone	work/ height (in ma)

As the 'meaning extension' explains itself, it is the frequent mechanism to express new concepts by using existing words, which then take on new meanings that are metaphorically related to their old meanings. The second column of the data in (54) refers to the basic meaning of a word; whereas the third column indicates the extended meaning of the respective word. The data in (54) clearly show that extension of meaning is one of the methods that are used during the preparation of primary school textbooks. Just like the standardization committee of Oromo, the textbook developers have used the human body parts as source in order to designate new concepts. For instance, to easily designate and teach the concept 'angle' in mathematics, the experts simply extended the original word *kofa* 'trouser' to 'angle'. Moreover, some of the lexical items stated above, are homonym in nature having several meanings. For instance, in the extended senses, the concept *hojjaa* 'height of someone' has multiple layer of meanings such as 'work and height (in Ma)'. This is not the only word to have more than one meaning. The concept *adda* 'forehead' in data (54) is another homonym word having several meanings like 'front' (military), political organization and 'luck'. Such concept is found to have multiple meaning in other languages such as Amharic as it is stated in Samule (2017:119). Generally, body parts of

human beings are very productive to designate new concepts in Oromo like that of Amharic and Hadiyya.

5.2.1.1.1.2 Meaning extension from animals' bodies and farming activities

In Oromo, designating new concepts via extending meaning from the existing words that are related to animals and farming activities is also a well known method to enrich the language. The following data are collected from grade 8 biology and social study text books that can clearly be an evidence for the above claim.

(55)	meaning 1	meaning 2
<i>beneensa</i>	hyena	he who do not satisfy easily
<i>bo'oo</i>	furrow (of a ploughing)	line (of paragraph)
<i>calla</i>	harvest	coins (smaller money notes)
<i>gaanfa</i>	antler	horn
<i>hantuta (tee)</i>	mouse	mouse (of computer)
<i>harree</i>	donkey	sweep/ he who did not understand easily
<i>hoolaa</i>	sheep	ingenous
<i>horii</i>	cattle	treasure
<i>leenca</i>	lion	brave
<i>saree</i>	dog	child (lovely child)
<i>tuqaa</i>	mole	full stop (period)

In addition to the lexical elaboration method via the extension of meaning from the body parts of a human being, the data that are collected from the text books clearly show that the name of the animals and farming activities are the sources for new concept designation. There are some base words that have homonym characteristics having more than two meanings. For instance, in Oromo the lexical items *horii* 'cattle' is extended beyond its original meaning to designate the concept 'money' and 'treasure'.

5.2.1.1.1.3 Extension from parts of plants and food

It is evident that the education sector has also played a great role in elaborating Oromo as several new concepts are designated through meaning extension from plants, parts of plants and various types of food. The data in (56) are collected from grade 6, 7 and 8 biology, social study and Oromo text books.

(56) Oromo	meaning 1	meaning 2
<i>cuunfaa</i>	juice	summary
<i>damee</i>	bough (of a plant)	branch (of a bank)
<i>damma</i>	honey	lovely
<i>dhangaa</i>	traditional food	nutrient
<i>hundee</i>	root of plant	root of word
<i>mi'aa</i>	sweat	costly
<i>sanyii</i>	seed	type (variety)/ ethnic group

As the data in (56) clearly show, plants, parts of plants and the food we eat can also be the sources for the designation of new concepts. The extended meanings in column three are extended from the basic meanings from column two. Some lexical item like *sanyii* 'seed' is homonym words in nature which are having an extended meaning (i.e., type /variety/ ethnic group).

5.2.1.1.1.4. From materials and shelter

There are also evidences to conclude that various materials which are being used by human beings that became the sources for the enrichment of Oromo. The data in (57) are collected from from grade 6, 7 and 8, social study, biology and Oromo text books.

57) Oromo	meaning 1	meaning 2
<i>gaachana</i>	shield	barricade
<i>gaalee</i>	creeper	phrase (grammar)
<i>gabatee</i>	material (to wash our leg)	table/ keyboard of a computer/numberplate
<i>geeba</i>	cup	penis
<i>mana/no</i>	house	digit (of maths)
<i>siiqqee</i>	stick	axis (maths)
<i>warqii/ee</i>	gold	important/good/ 'ʔi'nsət' in Amharic

As we can see in (57), the lexical gap of Oromo is seen to be enriched through taking the name various material and extending the meaning to designate another new concept. Some base words in the above data have more than two meanings. For instance, *gabatee* 'traditional material used to wash our leg or hand' and *warqii/ee* have multiple layer of meanings that is table/ keyboard of computer and important /good/ 'ʔi'nsət' (kind of plant) respectively (See also Griefenow-Mewis & Tamene 1995:79).

5.2.1.1.2 Redefinition of words

Redefinition of existing word is also a frequent mechanism to express new concepts in Oromo. This can be seen from the data illustrated below. The data in (58) are collected from grade 7 mathematics, grade 10 and 11 Oromo textbooks.

(58) Oromo	meaning 1	meaning 2
<i>bu'aa</i>	use	result/ profit (in business)
<i>cabsuu</i>	to brake	demoralize/ to lower the price of something
<i>dhaamuu</i>	inform	switch (of computer), off (of light)
<i>guduunfaa</i>	to tie	summary
<i>maxansuu</i>	belay (to stick)	print
<i>mirga</i>	right (direction)	a person's duty (in law)
<i>salphisu</i>	disrespect (abjection)	to simplify (in mathematics)
<i>tarkaanfii</i>	walk	measure/ step (mathematics)/ progress

As can be seen from (58), the basic meanings (general vocabulary) of the lexical items are given in the second column, whereas the specific field (subject) area meanings are given under the third column. New concepts in a particular area are being designated via redefinition of the words. For instance, the lexical item *salphisu* 'disrespect (abjection)' is redefined to designate specific concept called *simplify* in mathematics. Some other lexical items like *cabsuu* 'to break' and *tarkanfii* 'walk' have more than three meanings by redefining the base meanings. The former (i.e. *cabsuu*) lexical item is redefined to designate the concept 'demoralize' and other third meaning called 'to lower the price of something' whereas the later lexical item (i.e. *tarkaanfi* 'walk' is redefined to represent other subject area meaning (i.e. measure/step (mathematics)/ progress).

5.2.1.1.3 Derivation

The concept derivation can be defined as deriving something out of the existing word that creates more meaningful words from it (Refat 2015:17). When the derivationist noted derivation from the broader sense, it is the process of making new words out of the existing ones. We may consider the often cited English example in which the word *do* represent an open class word in the branch of verbs while adding suffix *-er* will create its new meaning with a new word that is *doer*. This affix changes its branch from verb to noun. Derivation may also create a new word via changing its meaning being in the same lexical category as in, for instance, *do--did-done*.

Returning back to the present study's concern, derivation is one of the methods of coining new words in Oromo like any other languages. The main objective of this section is not to discuss the morphology of the language in general. But, it is to describe how derivation is used by the textbook writers to cover the lexical gaps of the language. Derivational suffixes are added to the root or stem of the word so as to create a new word (see also Baye 1986:12 & Temesgen 1993:9). In this sense, Oromo is a suffixing language as we can see in the following section.

The lexicon of Oromo is enriched in the textbooks via various methods. Accordingly, various new words in the syntactic category of noun, adjective and verb are coined. Nominalization,

verbalization and adjectivalization are discussed as strategies of lexical elaboration of the language in the textbooks.

There are a number of noun-creating formatives in Oromo which are evident in the textbooks that this study consulted. Looking into the textbooks, one can note that nouns, verbs and adjectives are the main sources of coining new nouns. Such formations are described as follows.

5.2.1.1.3.1 Nominalization

The lexical elaboration via adjectival nominalization is one of the methods that the Oromo textbook compilers used to enrich the language. Oromo abstract nominals are derived from adjectives by the addition of different suffixes such as *-ina*, *-ummaa*, *-eenya* as shown in (59). The data are collected from grade ten, eleven and twelve Oromo textbooks.

(59) Base gloss	affix (NMZR)	derived nominals	gloss
<i>addii</i> 'white'	<i>-eenya</i>	<i>aaddeenya</i>	'whiteness'
<i>furdaa</i> 'fat'	<i>-ina</i>	<i>furdina</i>	'fatness'
<i>gamna</i> 'wise'	<i>-ummaa</i>	<i>gamnummaa</i>	'wisdom'
<i>goota</i> 'brave'	<i>-ummaa</i>	<i>gootummaa</i>	'bravery'
<i>jabaa</i> 'strong'	<i>-eenya</i>	<i>jabeenya</i>	'strength'
<i>salphaa</i> 'simple'	<i>-ina</i>	<i>salphina</i>	'ignominy'

As seen in the set of examples of derivation of nouns from adjectives in (59), abstract nominals are formed through attaching suffixes to the adjectives. This method is extensively applied by the textbook writers to enrich the lexicon of the language. Of the nominalizing suffixes stated in (59), *-ummaa* can appear both with noun and adjectives in the process of formation of nouns whereas *-eenya* and *-ina* appear with adjectives. These nominalizing suffixes are unpredictable in terms of their distribution as they sometimes complement and in others cases they substitute each other as in, for instance, *jabina* and *jabeenya* both to mean 'strength'. There are also cases in which they complement as in *addeenya* 'whiteness' unacceptable to write as *addina* (see also Baye 1986:120; Temesgen 1993:9).

New words in the syntactic category of nouns are also formed from other nouns. The resultant nouns that are formed by the combination of existing nouns and nominalizing suffixes are abstract nouns as seen in (60). The following data are taken from grade eight Oromo textbook.

(60) Noun	gloss	affix	noun	gloss
<i>addunyaa</i>	'world'	<i>-ummaa</i>	<i>addunyummaa</i>	'globalization'
<i>fira</i>	'relative'	<i>-ooma</i>	<i>firooma</i>	'affiliation'
<i>hiriyaa</i>	'friend'	<i>-ummaa</i>	<i>hiriyummaa</i>	'friendship'
<i>nama</i>	'man'	<i>ooma</i>	<i>namooma</i>	'humanity'

Here, it is possible to note that the suffix *-ummaa* is more productive than *-ooma* as it is used several times by the textbook compilers to coin a number of new words both from adjectives and other nouns. Still the distribution of these affixes is unpredictable, since they are not phonologically triggered.

The strategy of derivation of nouns from verbs is also extensively used by the textbook writer so as to enrich the vocabulary of the language. A number of suffixes are attached to the root of the verb to result in verbal nominal. The data in (61) show nominalizing suffixes, verbal roots and the resultant derived nominal. The data are collected from grade eleven Oromo textbook.

(61) Base	affix	derived nominals	gloss
<i>bar-</i> 'learn' (know)'	<i>-umsa</i>	<i>barumsa</i>	'education'
<i>dhug-</i> 'drink'	<i>-aatii</i>	<i>dhugaatii</i>	'beverage'
<i>gargaar-</i> 'help'	<i>-sa</i>	<i>gargaarsa</i>	'cooperation'
<i>haam-</i> 'mow'	<i>-aa</i>	<i>haamaa</i>	'a person who cut grass'
<i>kuf-</i> 'fall'	<i>aatii</i>	<i>kufaatii</i>	'collapse'
<i>mur-</i> 'cut'	<i>-too</i>	<i>murtoo</i>	'decision'
<i>qor-</i> 'examine'	<i>-maata</i>	<i>qormaata</i>	'examination' ('challenge')
<i>taajab-</i> 'observe'	<i>-ii</i>	<i>taajabbii</i>	'disapproval'

As can be noted from (61), there are a number of noun-creating form can be attached to verbal roots to result in nominals. Like any other nominalizing affixes we have seen in this subsection,

it is also difficult to predict the distribution of the above affixes with precision as they are not governed by any fixed rule.

The textbook compilers also used gerundive as a strategy of enriching the language. In such strategy, nominals are formed by the suffixing of the affix *-uu* to the verbal roots as shown in (62). The data under (62) are collected from grade eight Oromo textbook.

(62) Base	affix	derived nominals	gloss
<i>barrees-</i> 'write'	<i>-uu</i>	<i>barreessuu</i>	'writing'
<i>duubbis-</i> 'read'	<i>-uu</i>	<i>dubbisuu</i>	'reading'
<i>hat-</i> 'steal'	<i>-uu</i>	<i>hatuu</i>	'stealing'
<i>saam-</i> 'rob'	<i>-uu</i>	<i>saamuu</i>	'plundering'
<i>waaltes-</i> 'standardize'	<i>-uu</i>	<i>waaltessuu</i>	'standardizing'

The above gerundive nominals are resulted by attaching affixes to the verbal roots. Although the data are taken from one textbook, such strategy is extensively used throughout most of the textbooks that this study consulted.

Textbook compilers have also used agentive as a means of new word formation. Agentive nominals are also among the nominal that are the result of the combination of verbs roots and suffixes such as *-aa* and *-tuu*. Such strategies are evident in the textbooks. The data in (63) are collected from grade three and eight Oromo textbooks.

(63) Base	gloss	affix	derived nominals	gloss
<i>barbaad-</i>	'search'	<i>-aa</i>	<i>barbaadaa</i>	'searcher'
<i>gargaar-</i>	'help'	<i>-aa</i>	<i>gargaaraa</i>	'assistant/helper'
<i>haadh-</i>	'remove'	<i>-aa</i>	<i>haadhaa</i>	'cleaner'
<i>haammatt-</i>	'engulf oneself'	<i>-tuu</i>	<i>haammattuu</i>	'parenthesis'
<i>hacuuc-</i>	'oppress'	<i>-aa</i>	<i>hacuucaa</i>	'oppression'
<i>kut-</i>	'cut'	<i>-aa</i>	<i>kutaa</i>	'cutter (room) (grade)'
<i>madaal-</i>	'measure'	<i>-tuu</i>	<i>madaaltuu</i>	'scales'

The distribution of such suffixes is difficult to predict with precision. For instance, the affixes *tuu* and *aa* can be affixed to the same verbal root as in, for instance, *gargaaraa* 'assistant (masculine)' and *gargaartuu* (feminine)'. Here, the affixes *-aa* and *-tuu* can be noted from their behavior point of view. It seems problematic to treat these affixes as either derivational or inflectional as they are used both as derivation (they are attached to a verb root to coin new words) and as inflectional since they are suffixed to clearly distinguish gender i.e., *-aa* and *-tuu* as masculine and as a feminine marker respectively (see also Temesgen 1993:15).

In addition to the coinage of new nouns from other nouns and adjectives, nouns in Oromo can also be formed by derivation of the existing verbs. In his study on nominalization via verbal derivation of Amharic, Tigrinya and Oromo, Shimelis (2014) described the lexical derivation of noun from verbs. Shimelis's study on verbal noun formation of Oromo is a description of the phenomena as it appears in the spoken form of the language. The present study, however, examines, in this subsection, the verbal nominalization strategies of enriching Oromo as it appears in the written form of the language (particularly in the textbooks here in this subtopic).

Oromo textbook compilers also used the existing verbs so as to coin a new word (i.e., verbal noun). There are several such verbal nominals in the textbooks which are formed by adding the nominalizing morpheme *-uu* to the stem (verb). The following verbal nouns are collected from grade twelve Oromo textbook as an illustration.

(64) verbal nouns	gloss
<i>ibs-uu</i>	
describe-VNM	'describing'
<i>qood-uu</i>	
divide-VNM	'dividing'

Such nominalization strategies are used in all most all the textbooks this study consulted. This type verbal nominalization from non-derived simple verb is a very productive technique of enriching the language which is evident in the textbooks.

The textbook compilers of Oromo have also used non derived verbs as a source of coining new words. The education sector is also found to coin new words through the nominalization of non-derived composite verbs which include ideophone (the non inflecting constituent) and the auxiliary verb (the inflecting or changing verb that result into verbal noun). The following data is collected from grade 10, 11 and 12 Oromo textbooks.

(65)	Ideophone	root	verbal noun	causative	verbal noun
a)	<i>biliq</i>	<i>godh-</i>	<i>biliq godh-uu</i>	<i>biliq</i>	<i>goch-i-siis-uu</i>
	sudden kill	make s	udden kill make- VNM	sudden kill	make-ep-CS-VNM
			'killing suddenly'		'making (somebody) to kill suddenly'
b)	<i>qah jedh-</i>		<i>qah jech-uu [jedh-uu]</i>	<i>qah jech-i-siis-uu [jedh-i-siis-uu]</i>	
	dry say		dry say- VNM	dry say-ep-CS-VNM	
			'drying'		'causing something to become dry'

As can be seen from the data in (65) it is the verb that takes the nominal marker in order to change its verbal function into the noun function irrespective of with which (either the directional or the auxiliary) the verb combines.

In addition to the above forms of verbal nominalization (nominalization process of verbs that come with ideophones), there are also other nominals which are evident in the textbooks. The causative verbal nominalization includes the verbal root, affix and causative. The data for such assertion are taken from grade six Oromo textbook. In fact, such method of coining new words is found in almost all textbooks that this study consulted.

(66)Base	affix	causativized verb	affix	causative verbal nouns	gloss
<i>calaq</i> - 'reflect'	<i>-i-siis</i>	<i>calaqisiis-</i> 'cause to reflect'	<i>-uu</i>	<i>calaqqisiisuu</i>	'cause to reflection'
<i>dafq-</i> 'sweat'	<i>-i-siis</i>	<i>dafqisiis-</i> 'cause to sweat'	<i>-uu</i>	<i>dafqisiisuu</i>	'cause to sweat'
<i>madd-</i> 'well'	<i>-i-siis</i>	<i>maddisiis-</i> 'cause to well'	<i>-uu</i>	<i>maddisiisuu</i>	'cause to well'
<i>qorr-</i> 'cool'	<i>-i-siis</i>	<i>qorrisiis-</i> 'cause to cool'	<i>-uu</i>	<i>qorrisiisuu</i>	'cause to cooling'

Apart from the formation of verbal nominals (verbal nouns) from the non-derived verbs, deverbal nouns of Oromo can be coined from the roots of non-derived verbs. Nouns formed with such operation are deverbal nominals. The data are collected from grade six and twelve Oromo textbooks.

(67) Root	affix	deverbal noun
a) <i>qajeelch-</i> direct (guide)	<i>-a</i>	<i>qajeelch-a</i> direct- NMZR 'direction'
b) <i>qor-</i> examine	<i>-maata</i>	<i>qor-maata</i> examine- NMZR 'examination'
<i>hir-</i> divide	<i>-maata</i>	<i>hir-maata</i> divide-NMZR 'division'
c) <i>fil-at</i> choose-MD	<i>-noo</i>	<i>fil-at-noo[filannoo]</i> choose for oneself' choose- NMZR 'choice'

As the above example shows, the formation and meanings of Oromo derived nouns vary depending on the morpheme suffixed in to the root. Hence, there is an irregularity in the morphological pattern during the creation of deverbal noun.

Looking into all the nominalization processes which are attested in the textbooks, one can note that all verbal roots on which any nominalizing affixes attached are consonant final. To the contrary, the adjectives and nouns that are used as a base word for coining new lexicon are not roots as in verbal nominalization; but words. High tone is also applied in the process of creating new nouns (verbal nouns) from verbs (see also Shimelis 2014:62). On the other hand, the nominalizing affixes that we have seen under derivation do not be treated with any rule. The distribution of derivational affixes cannot be accounted with precision.

5.2.1.1.3.2 Lexical elaboration via verbalization

Since the main objective of subtopic is to describe the verbalization techniques used by the textbook writers to elaborate the lexicon of the Oromo, the study did not go into the details of the morphology of verbalization. The subtopic only describes how the textbook writers used the verbalization technique to elaborate the language (see 4.1.2.2). Apart from the lexical elaboration of Oromo via nominalization, there are verbalization and adjectivization. Textbook compilers applied causativizing suffixes which are - (i)s and -(i)sis and (i)siis- to enrich the lexicon of the language. Such verbalization process takes place through attaching causativizing suffixes -(i)s, -(i)sis and -(i)siis to verbal roots as seen in (68). The following data are taken from grade three Oromo textbook.

(68) Base	affix	causativized verb	gloss
<i>bar-</i> 'learn'	-siis-	<i>barsiis-</i>	'make learn'
<i>gees-</i> 'materialise'	-sis-	<i>geessis-</i>	'cause to materialise'
<i>jig-</i> 'tumble down'	-s-	<i>jigs-</i>	'cause to tumble down'
<i>mad-</i> 'gush'	-i-siis-	<i>maddisiis-</i>	'cause to gush'
<i>sirb-</i> 'sing'	-i-siis-	<i>sirbi-siis-</i>	'cause to sing'
<i>taphat-</i> 'play'	-is-	<i>taphatis-</i> [<i>taphachis-</i>]	'cause to play'
<i>tur-</i> 'linger (stay)'	-siis-	<i>tursiis-</i>	'cause to stay'

Here, one can note that verbal roots are among the sources for the verbalization of the language. Textbook compilers used such process so as to enrich the lexicon of the language. I partly agree with Baye's (1986:126) argument that distribution of causativizing suffixes is to some extent determined by the transitive intransitive feature of the verbs undergoing the process. Accordingly, intransitive verbs require -i-s- whereas verbs which are transitive take -i-sis- and -i-siis-. But this is not always the case as there are some verbs which are intransitive; but that take -siis- as in *tur* -siis- 'cause to stay'. Generally, -sis and -siis- follow vowel length dissimilation rule to the length of the vowel of the previous syllable.

The sources of coining new verbs are not only the existing verbs. Verbalization can also take place by the attachment of verbalizing suffixes to nouns and adjectives as in (69). The following illustrative data are taken from grade three Oromo textbook.

(69) Base	affix	causativized verb	gloss
<i>boca</i> 'shape'	<i>-is-</i>	<i>bocis-</i>	'cause to shape'
<i>firooma</i> 'affiliation'	<i>-s-</i>	<i>firooms-</i>	'match (affiliate)'
<i>furdaa</i> 'fat'	<i>-is-</i>	<i>furdis-</i>	'make fat'

As can be seen from (69), Oromo also enriched itself in the textbooks through internal means by suffixing verbalizing affixes to nouns and adjectives. The resultant of such process is causativized verb.

Apart from the verbalization process via causativization, textbook writers also coined new words through passivization. Here, transitive bases are the sources for the derived passive verbs. The following data are collected from grade four Oromo textbook.

70) Base	affix (passive)	derived forms	gloss
<i>sarar-</i> 'underline'	<i>-am-</i>	<i>sarar-am-</i>	'be underlined'
<i>godh-</i> 'make'	<i>-am-</i>	<i>godh-am-</i>	'be made'
<i>hayyam-</i> 'permit'	<i>-am-</i>	<i>hayyam-am-</i>	'be permitted'
<i>jedh-</i> 'say'	<i>-am-</i>	<i>jedh -am-</i>	'be said'
<i>kabaj-</i> 'respect'	<i>-am-</i>	<i>kabaj-am-</i>	'be respected'
<i>mur-</i> 'cut'	<i>-am-</i>	<i>mur-am-</i>	'be cut'

Passivization, as we can see in data (70), is one of the productive means of elaborating the lexicon of the language. Such strategies of enrichment are also used by the textbook compilers to cover the lexical gaps. Passive verbs in Oromo are derived from transitive bases by the addition of the suffix *-am-*.

There are also other strategies, other than causativization and passivization, which the Oromo textbook compilers used to coin new lexicon (verb). The language is found to enrich its lexicon through the formation of middle verbs. The following illustrative data are taken from grade one and two Oromo textbooks.

(71)	Base	affix (MD)	derived	gloss
	<i>baq-</i> 'melt'	<i>-at-</i>	<i>baq-at-</i>	'melt for oneself'
	<i>dabal-</i> 'add'	<i>-at-</i>	<i>dabal-at-</i>	'add for oneself'
	<i>dhaabb-</i> 'plant'	<i>-at-</i>	<i>dhaabb-at-</i>	'plant for oneself'
	<i>dhageeff-</i> 'listen'	<i>-at-</i>	<i>dhageeff-at-</i>	'listen for oneself'
	<i>dhal-</i> 'bring forth'	<i>-at-</i>	<i>dhal-at-</i>	'born'
	<i>jaal-</i> 'love'	<i>-at-</i>	<i>jaal-at-</i>	'love for oneself'

Looking into the data, one can note that the process of middle formation is non-category changing as it derives middle verbs from other verbal base. Hence, the construction of middle through the addition of the suffix *-at-* to existing verbal base is one of the strategies by which Oromo textbook compilers used to elaborate the lexicon of the language. On the other hand, stative verbs can be derived from adjectivals and nominals with the suffix *-at-*. The following stative verbs are collected from grade five social study and Oromo textbooks.

(72)	Base	affix (MD)	derived statives	gloss
	<i>dheeraa</i> 'long'	<i>-at-</i>	<i>dheer-at-</i>	'become long'
	<i>hamaa</i> 'bad'	<i>-at-</i>	<i>hamm-at-</i>	'become bad'
	<i>jabaa</i> 'strong'	<i>-at-</i>	<i>jab-at-</i>	'become strong'
	<i>onnee</i> 'heart'	<i>-at-</i>	<i>onn-at-</i>	'become brave'

In the process of the formation of stative verbs, Oromo textbook compilers used various lexical categories such as noun and adjectives, as seen in (72), to develop the lexicon of the language. Here, we can note that unlike the passivization process, in which there is no categorical change, the process formation of stative verbs is a category changing event.

5.2.1.1.3.3. Adjectivization

Adjectivization is a process by which new words in the lexical category of adjectives are formed through the addition of suffixes to the existing other words. Adjectivization is one of the strategies with which the Oromo textbook compilers developed the lexicon of the language. Adjectivization existed in almost all of the textbooks that this study consulted. The following illustrative data are taken from grade five social study textbook.

(73) Base word	affix	derived forms	gloss
<i>ammaayy</i> - 'be modern'	-aa/ <i>tuu</i>	<i>ammaayy-aa/tuu</i>	'modern (masc/ fem)'
<i>diim</i> - 'be red'	-aa <i>tuu</i>	<i>diim-aa/tuu</i>	'red (masc/fem)'
<i>diriir</i> - 'be flat'	-aa/ <i>tuu</i>	<i>diriir-aa/tuu</i>	'flat (masc/ fem)'
<i>jab</i> - 'be strong'	-aa/ <i>tuu</i>	<i>jab-aa/tuu[duu]</i>	'strong (masc/fem)'

Although there are other adjectivizing suffixes, as can be seen in the examples to come, the suffixes *-aa* and *-tuu* are the most productive adjectivising suffixes that are evident in the corpus of textbooks that this study consulted. The controversial issue, here, is how to, clearly, determine the lexical categories of the base words. One can categorize the base words as adjectivals since they have inherent features of adjectives and affixes *-aa* and *-tuu* can be considered as inflectional affixes representing masculine and feminine respectively. For instance, the base *jab*- can be regarded an adjective and the affixes *-aa/tuu* as an inflectional morpheme indicating masculine and feminine gender as in *jab-aa/ jab-tuu [duu]* 'strong (masc./fem)'. On the other hand, the same base words/root can be considered as verbal root as each of the above roots can take the verbalizing suffixes such as *-at*, *-ess* and others so as to result in derived verbs. Adjectives are also found to be derived in the textbooks from the verbal roots with the addition of suffixes *-aa* and *-tuu* to the bases as seen in (74). The data are taken from grade five social study textbook.

(74) Base	affix	derived forms	gloss
<i>badhaadh</i> - 'prosper'	-aa/ <i>tuu</i>	<i>badhaadh-aa/tuu</i>	'prosperous (masc/fem)'
<i>marmaar</i> - 'horrify'	-aa/ <i>tuu</i>	<i>marmaar-aa/tuu</i>	'horrible (masc/fem)'
<i>mirkaneess</i> - 'corroborate'	-aa/ <i>tuu</i>	<i>mirkaness-aa/tuu</i>	'corroborative (masc/fem)'
<i>murteess</i> - 'decide'	-aa/ <i>tuu</i>	<i>murteess-aa/tuu</i>	'desicive (masc/fem)'

Unlike the bases (root) in data (72) and (73), in (74), the roots on which the suffixes attached do not have an inherent feature of adjectives. Rather, these roots are verbal roots and become sources for adjectivization as a means of elaboration of the lexicon of the language in the textbooks.

Though not as productive as the adjectivizing suffixes such as *-aa* and *-tuu*, there are also other adjectivizing suffixes. The following illustrative examples of adjectivization are taken from grade five social science and grade eight biology textbooks.

(75) Base	affix	derived forms	gloss
<i>dheed-</i> 'pasture'	<i>-ii</i>	<i>dheed-ii</i>	'raw'
<i>guut-</i> 'fill'	<i>-uu</i>	<i>guut-uu</i>	'full'
<i>kennam-uu</i>	<i>-uu</i>	<i>kennam-uu</i>	'gifted'
<i>sob-</i> 'lie'	<i>-oo</i>	<i>sob-oo</i>	'false (of false)'
<i>qab-</i> 'possess'	<i>-eessa</i>	<i>qab-eessa</i>	'that possess (wealthy)'
<i>qar-</i> 'sharpening'	<i>-oo</i>	<i>qaroo</i>	'sharp (eye)'

As can be seen from the illustrative examples above, the resultant derived forms possess all the features of adjective as, for instance, like any other adjectives; they can appear after the word *baay'ee* 'very'. Looking into all the adjectivization processes that have been discussed so far, one can note that it is the verbal root which are more productive than other the roots of lexical categories. All in all, regarding the adjectivization as a means of lexical elaboration, this study argues and supports Temesgen's (1993:33) argument that, of the existing suffixes, *-aa* and *-tuu* are the more productive adjectivising suffixes. These suffixes, however, are at the same time, serves to indicate the gender of the derived form.

5.2.1.1.4 Compounding in the textbooks

The distinction between compounds and multiword-expressions (phrasal words) and in some of the cases even phrases is problematic. Such problems arise due to several reasons. One could be due to the fact that there are theoretical problem of identifying a sharp divide among compounds, multiword expressions and phrases. Another one is there is no consensus over the criteria for definitions and distinctions of these notions and as a result authors define and make a demarcation among these notions in their own ways. Not all authors also differentiate these concepts in various languages. Moreover, there is no universal criterion to distinguish compounds, multiword expression and proper phrase that works cross-linguistically. Stated by Masini (2009:268), there is a considerable debate in Italian linguistics regarding the above

concepts. Masini (2009:268) states that from the construction grammar perspective, all the structures of the language are regarded as constructions and what vary is their size and internal complexity, their function, and their position within the construction. The author points out that in principle we can have constructions that share structural properties and constructions that share semantic properties and this happens with phrasal lexemes, phrases and morphological words or, better, lexemes including compounds. The author concludes that morphological lexemes and phrasal lexemes may share the function without sharing the structure; phrasal lexemes and phrases proper may share the structure but not the function (Masini 2009:268). Compounds are by definition morphological lexemes, even though they are certainly more phrase-like than other morphological objects due to their internal composition (Masini 2009:268). With almost the same outlooks Ziering and Van der Plas (2014:15) consider compounds to constitute a continuum from a fully compositional to an idiosyncratic word formation and that compounds are found at the interface between words and phrases (syntax).

There are problem of identifying a sharp divide among compounds, multiword expressions and phrases in Oromo too. Therefore, this area needs further investigations. Some defining features of Oromo compounds that distinguish them from Oromo syntactic objects are, however, discussed under section 4.1.2.4 with non exhaustive. Coming back to our main concern, compounding is used as a means of enriching Oromo in various textbooks so as to narrow the existing lexical gaps. The following section describes compounds, features of compounds that are being observed in the textbooks. Compounding is one of the most productive word formation types which are observed in various primary textbooks and secondary Oromo textbooks which will be discussed in the following subtopics.

5.2.1.1.4.1 Types of compounds of Oromo from orthographic perspective

From the perspective of their orthography, Oromo compounds can be closed, hyphenated or open. Three of the compounding strategies have been used as a means of enriching the lexical gaps of the language. Each of these forms is presented and discussed in the following sections.

5.2.1.1.4.1.1. Closed compounds

Oromo closed compounds are formed from two or more which are written (spelled as a single word without having any space or hyphen in between or among them. The data collected from grade 8 biology and from grade 10 and 11 Oromo textbooks show the closed compounds as it is depicted in (76). The following data clearly show that the closed compounding method has been used during the compilation of Oromo textbooks.

(76)	base word	compounds	gloss
(a)	<i>caala</i> (A) 'greater' + <i>baasii</i> (N) 'expenditure'	→ <i>caalbaasii</i> (N)	'auction'
(b)	<i>caasa</i> (N) 'structure' + <i>luga</i> (N) 'language'	→ <i>caasluga</i> (N)	'grammar'
(c)	<i>faalla</i> (A) 'opposit' + <i>darbeeyyii</i> (A) 'transitive'	→ <i>faaldarbeeyyii</i> (A)	'intransitive'
	<i>gocha</i> (N) 'action' + <i>ibsa</i> (N) 'description'	→ <i>gochibsa</i> (N)	'adverb'
	<i>maqaa</i> (N) 'name' + <i>ibsa</i> (N) 'description'	→ <i>maqibsa</i> (N)	'adjective'
(d)	<i>saree</i> (N) 'dog' + <i>diida</i> (N) 'outside'	→ <i>sardiida</i> (N)	'fox'
(e)	<i>nama</i> (N) 'human' + <i>tolche</i> (V) 'made'	→ <i>namtolchee</i> (A)	'man-made'
(f)	<i>seelii</i> (N) 'cell' + <i>qeenxee</i> (A) 'single'	→ <i>seelqeenxee</i> (A)	'unicell'

The data in (76) show that the textbook developers have used compounding as a means of filling the lexical gaps of the language. As seen from the data in (76), all compounds are formed via the deletion of the final vowels of the first constituent. It can also be noted that all of the compounds are compounds that can be classified as closed compounds as they are spelled as a single word. The above data can also be noted from the lexical category of the exocentricity point of view. By looking at the data in (76e), it can be noted that Oromo compounds can have an absolute categorial exocentricity structure. This happens when the output is completely different from the input categories as in (76e) that hold $[N+ V]_A$ structure. Categorically, except the data in (76e), all of the compounds are endocentric. By the same criterion (lexical category), compounds under (76a and 76c) are right-headed. For instance, as seen from (76a), categorically, the head, *baasii*, is the constituent that imposes its categorial features to the whole compound. To the contrary, compounds which are shown in (76b, 76d and 76f) are left headed by the similar judgement. Here, one can also note that the majority of the compounds in (76) are attributive compounds such that the compound's constituents are linked by a modified-modifier relation. From the

combination point of view, structure of compounds of Oromo in (76) can be represented by [A + N]_N, [N + N]_N [A + A]_A, [N + V]_A and [N + A]_A. One has to note here that these patterns are only as per the above closed Oromo compounds. There might also be other patterns.

One can also see the data in (76) from the perspective of semantic combination. Accordingly, all closed compounds found under (76) are semantically endocentric as the semantics of the input constituents is similar to the semantics of the output. For instance, in *seelii* 'cell' + *qeenxee* 'single' → *seelqeenxee* 'unicell', the two constituents are semantically similar to the output. In other words, they have synonymic relationships. Here, absolute semantic exocentricity has not been observed in data (76). This does not mean that all Oromo close compounds are semantically endocentric. On the other hand, Oromo closed compounds can be categorically exocentric while being semantically edocentric as in *nama* (N) 'human' + *tolche* (V) 'made' → *namtolchee* (A) 'man-made'. To the contrary, there are cases in which categorical and semantic endocentricity overlaps as in *caasa* (N) 'structure' + *luga* (N) 'language' → *caasluga* (N) 'grammar'.

5.2.1.1.4.1.2 Open compounds.

There are also Oromo compounds which are formed via combination of two or more words in which the compounds are spelled with whites pace-separated words. There is evidence that can help us to conclude that Oromo textbook developers have applied the open compounding strategy to enrich the lexicon of the language. The following data are collected from grade 8 biology and grade 10 and 11 Oromo textbooks which can clearly show the stated fact.

(77)base words	compounds	gloss
(a) <i>aantima</i> 'object' (N) + <i>kallattii</i> 'direct' (A)	→ <i>aantima kallattii</i> (N)	'direct object'
(b) <i>adeemsa</i> (N) 'process' + <i>tirataa</i> (A) 'tardy'	→ <i>adeemsa tirataa</i> (N)	'evolution process'
<i>baruu</i> 'learn' (V) + <i>barsiisuu</i> 'teaching' (N)	→ <i>baruu barsiisuu</i> (N)	'teaching learning'
(c) <i>gochima</i> 'verb' (N) + <i>darbeeyyii</i> (A) 'transitive'	→ <i>gochima darbeeyyii</i> (N)	'transitive verb'
<i>gochima</i> (N) 'verb' + <i>faaldarbeeyyii</i> (A) 'intransitive'	→ <i>gochima faaldarbeeyyii</i> (N)	'intransitive verb'
<i>gochima</i> (N) 'verb' + <i>gocha</i> (N) 'action'	→ <i>gochima gochaa</i> (N)	'action verb'

(d) <i>irra</i> (PREP) 'on' + <i>deebii</i> (Av) 'again' → <i>irra deebii</i> (N)	'revision'
(e) <i>jecha</i> 'word' (N) + <i>diigala</i> (N) 'compound' → <i>jecha diigala</i> (N)	'compound word'
<i>mataa</i> (N) 'head' + <i>gaalee</i> (N) 'phrase' → <i>mataa gaalee</i> (N)	'head word'
<i>madda</i> (N) 'source' + <i>odeeffannoo</i> (N) 'information' → <i>madda odeeffannoo</i> (N) 'source'	
(f) <i>xinqooqa</i> (N) 'linguistics' + <i>sammuu</i> (N) 'brain' → <i>xinqooqa samuu</i> (N)	'psycholinguists'

As shown in (77), Oromo is found to enrich itself via compounding that result in open compound. The data in (77) clearly show that more open compounding is used in Oromo textbooks as a means of filling language area related lexical gaps than scientific terminology. Just like other types of compounds, Oromo open compounds could be edocentric or exocentric categorically and semantically. Categorically, except the data in (77d), in which absolute categorial exocentricity occurred ([PREP + A]_N) and the constituent in the head position does not impose its categorial features on the whole construction, all open compounds under (77) are endocentric.

Hence, one can summarize the categorial structure of the endocentric open compounds as [N + A]_N, [V + N]_N and [N + N]_N. In these cases, the head constituents transmit their category to their respective compounds. Among these structures [N + N]_N is the most productive form of compounding. When the open compounds in (77) are seen from categorial head points of view, all most all of them are left headed compounds. Most of the data in (77) are attributive open compounds having linked by a modified-modifier relation whereas the compounds like *baruu barsiisuu* 'teaching learning' and *irra deebii* 'revision' are subordinative compounds. In such cases, the two constituents (*irra* 'on' and *deebii* as well as *baruu barsiisuu*) are connected by a predicator-argument relation. In other words, these two constituents have an implicit dependence relationship. But, open compound constructions like the case of *galmee* 'record' (N) + *jechootaa* (N) 'of words' → *galmee jechootaa* (N) 'dictionary' is an attributive compound.

Textbook developers have also used complex compounding method for filling their lexical gap. For instance, the data in (77a), (77c) and 77f are complex open compounds. These compounds are made up of three and more words in which the first (in a few of the case, the first constituent being a blend and the second being compound itself) constituent being a blend and the second

constituent being a word. The open compound *xinqooqa samuu* 'psycholinguistics' can be a clear instance for the formation of compounds from the blend *xinqooqa* 'linguistics' and another word *samuu* 'brain'. The complex compound *gochima faaldarbeeyyii* 'intransitive verb' is even more complex than the previous which is formed by combining the blend *gochima* 'verb' and the closed compound *faaldarbeeyyii* 'intransitive'.

On the other hand, open compounds that we have seen in (77) can also be viewed from semantic approach. Accordingly, all open compounds under (77) are semantically endocentric as the semantics of the input constituents is similar to the semantics of the output. Categorical endocentricity and semantically exocentricity may or may not match in a language. In Oromo, however, the semantic and categorical endocentricity sometimes may not match. For instance, the open compound *irra deebii* 'revision' is categorically exocentric as we can see in (77d) whereas the same compound is semantically endocentric since the semantics of the input constituents is similar to the semantics of the output.

5.2.1.1.4.1.3. Hyphenated compounds

In addition to those closed and open, the education sector seems to extensively use hyphenated compounding techniques to enrich the lexical gaps that existed in the education area. The following hyphenated compounds which are collected from grade 8 biology and grade 11 Oromo textbooks are some of the evidences which show lexical elaborations while preparing textbooks.

(78)	base word		compounds	gloss
	<i>bu'aa</i> (N) 'use'	+	<i>qabeessa</i> (A) 'having'	→ <i>bu'a-qabeessa</i> (A) 'useful'
	<i>gadii</i> (PREP) 'down'	+	<i>lakkisuu</i> (V) 'drop'	→ <i>gad-lakkisuu</i> (V) 'to drop'
	<i>gidduu</i> (N) 'middle'	+	<i>galeessa</i> (A) 'that inter'	→ <i>giddu-galeessa</i> (A) 'average'
	<i>lubbuu</i> (N) 'soul'	+	<i>qabeeyyii</i> (A) 'that possess'	→ <i>lubbu-qabeeyyii</i> (A) 'living (alive)'
	<i>lubbuu</i> (N) 'soul'	+	<i>maleeyyii</i> (A) 'that does not possess'	→ <i>lubbu-maleeyyii</i> (A) 'non-living'
	<i>mataa</i> (N) 'head'	+	<i>hima</i> (N) 'sentence'	+
	<i>mul'istuu</i> (N) 'indicator'	→	<i>matim-mul'istuu</i> (N)	'subject indicator'
	<i>miila</i> (N) 'leg'	+	<i>soboo</i> (A) 'of false'	→ <i>miil-soboo</i> (N) 'false feet'
	<i>saala</i> (N) 'sex'	+	<i>malee</i> (A) 'without'	→ <i>saal-malee</i> (A) 'non sexual'

As seen from the data in (78), all the hyphenated compounds are formed through two processes. During the formation, the last vowel of the first constituent is deleted and then hyphenated to the second constituent without deletion of any of the segments of the second constituent. It has to be also noted that just like the formation of some of the open compounds which we have seen in (77a), (77c) and (77f), there are also hyphenated compounds in Oromo which are formed from a blend and from another words as in,for instance, *mataa* 'head' + *hima* 'sentence' + *mul'istuu* 'that show' → *matim-mul'istuu* 'subject indicative'. All hyphenated compounds stated under (78) are attributive compounds and most of them are, categorically, exocentric compounds. When viewed from semantics, they are all endocentric. On the other hand, there are problems as to how compounds should be written. In some of the cases a compound is spelled as a single word, sometimes as and open and still in other cases the same compound is written hyphenated. Such variations are presented and discussed in spelling variation section (see section 5.2.1.2). A quick look at the data in (78) reveals that the structures of the hyphenated compounds can be represented $[N+N]_N$, $[PREP + V]_V$, $[N+ A]_N$ and $[N+A]_A$.

5.2.1.1.4.2 Types of compounds of Oromo from categorical point view

The education sector of Oromo, particularly the textbooks has extensively enriched the lexicon of the language through compounding. The compounds which are collected from textbooks are categorically classified and discussed in the following sections.

5.2.1.1.4.2.1 Nominal compounds

In this subsection of lexical elaboration, the nominal compounding methods that are used as a means of enriching the language in the primary school textbooks are discussed. Nominal compounding is a common phenomenon of Oromo as it clearly evident that it is used by the standardization committee (as we have seen in the previous chapter) and by education expertise to cover the lexical gap of the language which we shall see in the following subsection. $N+N$, $N+A$, $V+ N$, $N + AP$ are among nominal compounding methods that are attested in various primary school Oromo textbooks as a means of lexical elaboration. The following data which are collected from the current grade eight mathematics, civic and ethical education, biology and chemistry textbooks of primary school clearly attest that the nominal compounds of the form $N +$

N→nominal compounds are the most frequently used compounding method during the preparation of the textbooks.

(79)	Base words		nominal compounds	gloss
a)	<i>abbaa</i> (N) 'father'	+	<i>sangaa</i> (N) 'cattle' → <i>abbaa sangaa</i> (N)	'anthrax'
	<i>sabbata</i> (N) 'girdle'	+	<i>waaqayyoo</i> (N) 'God' → <i>sabbata waaqayyoo</i> (N)	'rainbow'
(b)	<i>afaala</i> (N) 'floor'	+	<i>daree</i> (N) 'class' → <i>afaala daree</i> (N)	'floor'
	<i>annisaa</i> (N) 'energy'	+	<i>aduu</i> (N) 'sun' → <i>annisaa aduu</i> (N)	'solar energy'
	<i>bulbula</i> 'solution'	+	<i>ayoodinii</i> (N) 'iodine' → <i>bulbulaa yoodinii</i> (N)	'iodine solution'
	<i>buufata</i> (N) 'station'	+	<i>boba 'aa</i> (N) 'gas' → <i>buufata boba'aa</i> (N)	'gas station'
	<i>daaraa</i> (N) 'ash'	+	<i>soodaa</i> (N) 'sodium' → <i>daaraa soodaa</i> (N)	'sodium carbonate'
	<i>galmee</i> (N) 'record'	+	<i>jechoota</i> (N) 'words' → <i>galmee jechootaa</i> (N)	'dictionary'
	<i>haala</i> (N) 'manner'	+	<i>raawwii</i> (N) 'accomplishment' → <i>haala raawwii</i> (N)	'aspect'
	<i>mana</i> (N) 'house'	+	<i>murtii</i> (N) 'decision' → <i>mana murtii</i> (N)	'court'

As can be seen from the data in (79), Oromo is found to enrich its lexicons via nominal compounding. Such kind of compounding is very productive and prominent in the primary school textbooks that are consulted in the present study. All of the data stated under (79) are endocentric from the view of syntactic category. The majority of the above nominal compounds are also endocentric semantically as the meaning of the resultant compounds are derived from the meaning of their respective constituents. There are also Oromo nominal compounds which are semantically exocentric whereas categorically endocentric as in (79a). The nominal compound *abbaa sangaa* 'disease of cattle', however, is semantically exocentric as the semantics of the compound does not emanate from the meanings of its constituent (i.e. *abbaa* 'father'+ *sangaa* 'cattle') nor the compound has semantic head. This example clearly shows that there are Oromo nominal compounds that do not have semantic heads (being semantically exocentric) but that have categorical heads (being categorically endocentric). Semantically, the majorities of the nominal compounds are single and left headed compounds.

Most of the nominal compounds in (79) are attributive compounds in which the compound's constituents are linked by a modified-modifier relation. For instance, in *annisaa aduu* 'solar energy', *annisaa* 'energy' is specified by *aduu* 'solar' to particularly specify the kind of energy. Nominal compounds like *mata duree* 'topic' is a two headed coordinate compound which is characterized by a coordinating relation whose constituents are tied by the conjunction «and» implicitly. The two headedness of such a compound is from the semantic point of view. Except the last nominal compound, all the constituents of nominal compounds under (79) are written separated by white space, whereas the last nominal compound *sablammii* 'citizen' is written as a single word.

When noted from construction perspectives, most of the data in (79) seem to be a genitive phrase construction. However, their semantic clearly indicate that they are a lexicalized compound as it can be seen, for instance, in *galemee jechootaa* 'dictionary' *sabbata waaqayyoo* 'rainbow', *bulbula ayoodinii* 'iodine solution'. In this regard, Shimelis (2014:213) noted that the genitives and the compounds that we have seen in (79) are similar in construction. Unless we classify them semantically, we cannot formally identify them from each other by their construction process. This, however, does not mean that all Oromo nominal compound simply show a genitive relationship as there are nominal compounds where there is no genitive relation as in *jecha diigala* 'compound word'

The combination of noun and derived noun can also yield in nominal compounds in Oromo. Such combination, however, does not seem to be productive in Oromo at least in the materials that this study consulted. For instance, a nominal compound *gara jabeenya* 'cruelty' is constructed by combining *gaaa* 'stomach' (an existing noun) and *jabeenya* 'strength' (a derived noun).

Nominal compounds are also found to be formed from numerals and nouns during the preparation of primary school Oromo textbooks. Such kind of compounding is evident in grade 8 biology and mathematics as well as grade 6 and 9 Oromo textbooks shown in (80).

(80) Noun + numeral combination	nominal compounds	gloss
a) <i>bara</i> 'year' + <i>kuma</i> 'thousand' →	<i>bar-kumee</i>	'millennium'
b) <i>gama</i> 'side' + <i>tokko</i> 'one' →	<i>gamtokkee</i>	'single side'
c) <i>roga</i> 'side' + <i>baay'ee</i> 'many' →	<i>rog baay'ee</i>	'polygon'
d) <i>roga</i> 'side' + <i>sadii</i> →	<i>rog-sadee</i>	'triangle'
<i>roga</i> 'side' + <i>afur</i> 'four' →	<i>rog-afree</i>	'quadrilateral'
e) <i>seelii</i> 'cell [English]' + <i>heedduu</i> 'many' →	<i>seelheedduu</i>	'multicell'
<i>seelii</i> 'cell' + <i>qeenxee</i> →	<i>seelqeenxee</i>	'single cell'
f) <i>tarmii</i> 'term' [English] + <i>lamee</i> 'double' →	<i>tarmi lamee</i>	'binomial'
<i>tarmii</i> 'term' + <i>tokkee</i> 'single' →	<i>tarmi tokkee</i>	'monomial'

As presented in (80), the numeral + noun are also a productive means of nominal compound formation in Oromo. Technical terms in Oromo are found to be designated via the combination of numerals and nouns. Such means of elaborating the lexical gaps of the language are clearly evident in various Oromo textbooks consulted in the present study. One can note the data in (80) from the perspective of technique of nominal compound formation and the spelling of the resultant nominal compound. Although the end results of all of the combinations of numerals and nouns of the data in (80) are similar (i.e., the output is nominal compound), the process is varies. For instance, the nominal compounds in (80a) and (80d) are formed by leaving out the terminal vowel of the first constituent and hyphenating it to the second constituent. The same process of elision of the terminal vowel of the first constituent has taken place in (80b) and (80e); however, the constituents are merged together without any space and hyphen. In this case, the nominal compounds are spelled as single words. In (80c) and (80f), the same strategy of vowel elision of the first constituents is applied; but, the constituents are neither hyphenated nor merged after the elision of the terminal vowel of the first constituents. The resultant compounds are written with white space in between them.

On the other hand, all nominal compounds in (80) are attributive compounds having modified-modifier relationship. In the noun + numeral construction like in (80), the role of the second constituent (numerals) is to make the meaning of the first constituents (nouns) more specific. For

instance, in *bar-kumee* 'millennium', the semantic head *bara* 'year' is specified by the modifier *kuma* 'thousand'. All of the nominal compounds in (80) are semantically endocentric and are left headed both categorically and semantically. There is, however, inconsistencies (variations) as to how compounds are written. The same compounds are seen to be spelled more than three different ways in the various textbooks which this study consulted. Such variations are discussed in subsections to come.

Nominal compounds are also formed during the preparation of textbooks via combinations of adjective and nouns. The following are some of nominal compounds which are collected from grade 8 mathematics and civic and ethical education as well as from grade 9, 10 and 11 Oromo textbooks.

81 Noun + adjective	nominal compound	gloss
a) <i>abbaa</i> 'father' + <i>booraa</i> 'gray'	→ <i>abbaa booraa</i>	'owner of gray horse'
b) <i>baha</i> 'east' + <i>fagoo</i> 'far'	→ <i>baha fagoo</i>	'far east'
c) <i>buufata</i> 'station' + <i>xiqqaa</i> 'small'	→ <i>buufata xiqqaa</i>	'semi colon'
<i>dhugaatii</i> 'drink' + <i>laafaa</i> 'soft'	→ <i>dhugaatii laafaa (lalaafaa)</i>	'soft drink'
<i>gabatee</i> 'board' + <i>gurraacha</i> 'black'	→ <i>gabatee gurraacha</i>	'black board'
<i>hima</i> 'sentence' + <i>dachaa</i> 'double'	→ <i>hima dachaa</i>	'compound sentence'
<i>konkolaataa</i> 'car' + <i>xiqqaa</i> 'small'	→ <i>konkolaataa xiqqaa</i>	'minibus'
<i>miidhama</i> 'impairment' + <i>qaroo</i> 'visual'	→ <i>miidhama qaroo</i>	'visual impairment'
<i>sanbata</i> 'sunday' + <i>guddaa</i> 'big'	→ <i>sanbata guddaa</i>	'Sunday'

As can be seen from the data in (81), Oromo textbook developers applied noun + adjective combinations so as to form nominal compounds. This kind of lexical elaboration is evident in various primary and secondary Oromo textbooks which are consulted in the study. Looking into the data in (81), one can note that, except in (81a), all the resultant nominal compounds are semantically endocentric as their meanings are derived from the meanings of their respective constituents. In the case of data in (81a), however, is exocentric since the semantics of the resultant nominal compound does not emanate from the meaning of its constituents (i.e., the

compound does not have semantic head). The nominal compounds in (81b) and (81c) are left headed semantically. But, categorically, all nominal compounds under (81) are endocentric as the respective constituents in the head position impose their categorical features on the whole construction. From the orthography point of view, except the nominal compound under (81b) in which it is a closed compound, all the rest nominal compounds are open compounds. There are also some concepts like 'soft drink' which can be represented in both compound and blend as in *dhugaatii laafaa* and *lalaafaa* respectively. Here, the blend *lalaafaa* 'soft drink' is formed by the combination of itself (i.e. *laafaa* 'soft' + *laafaa* 'soft'). Such blend is formed with the strategy of taking the terminal consonant and vowel (cv) of the first constituent and attaching it to the initial part of the second constituent. Oromo nominal compounds can also be formed through the combination of a verb and a noun, although such combinations are rare (i.e., not as productive as other ways of combinations). The following data are collected from grade nine Oromo textbooks.

(82) Verb	noun	nominal compounds	gloss
<i>dafqee</i> 'toil'	+ <i>bulaa</i>	'he who live' → <i>dafqee bulaa</i>	'proletarian'
<i>horsiisee</i> 'breeding'	+ <i>bulaa</i>	'he who live' → <i>horsiisee bulaa</i>	'breeder'
<i>qotee</i> 'he who plough'	+ <i>bulaa</i>	'he who live' → <i>qotee bulaa</i>	'farmer'

Some of the nominal compounds that are formed by the combination of verb and nouns are now seen to be used parallel with other nearly expressions. For instance, the concept 'farmer' is seen to be written as both *qotee bulaa* and as *qonnaan bulaa* in various textbooks that this study consulted.

The combination of adposition and noun is another possible strategy of forming nominal compounds in Oromo. In such combination, adpositions usually appear first. The data in (83) are collected from grade six, nine, eleven and twelve Oromo and grade eight civics and ethical education textbooks.

(83) Adposion + noun	nominal compound	gloss
a) <i>dura</i> 'front' + <i>taa'aa</i> 'sit'	→ <i>dura taa'aa</i>	'chairman'
<i>gad</i> 'down' + <i>qabaa</i> 'having'	→ <i>gad qabaa</i>	'oppression'

<i>irra</i> 'on' + <i>butaa</i> 'he who pull' → <i>irra butaa</i>	'cluster of different consonant'
<i>irra</i> 'on' + <i>dhaaba</i> 'station' → <i>irra dhaaba</i>	'a surprise marriage'
b) <i>irra</i> 'on' + <i>deebii</i> 'returning' → <i>irra deebii</i>	'revision'
<i>keessa</i> 'in' + <i>deebii</i> 'returning' → <i>keessa deebii</i>	'revision'

As seen from (83), textbook compilers enriched the language so as to cover existing lexical gaps. The nominal compounds shown in (83a) are semantically exocentric as these compounds are compounds in which the semantics of the input constituents are very different from the semantics of the outputs. When we see the same compounds from the categorical point of view, they are all endocentric. To the contrary, the compounds in (83b) are endocentric both semantically and categorically. In most of the cases, the nominal compounds of Oromo which are formed from adpositions and nouns, like we have seen in (83), are open compounds.

¶

Another possible, but very rare instances of nominal compound formation is adjective and noun combination. Temesgen (1993:34) stated that the existence of such nominal compound formation though he did not explain whether this combination is productive or not. But, the current study claims that such formation is very few. In the nominal compounds like *cimaa* 'clever' + *barataa* 'student' → *cimaa barataa* 'clever student' and *gurraacha* 'black' + *cilee* 'charcoal' → *gurraacha cilee* 'charcoal-black', the second constituent of the nominal compound qualify the first constituent of compound. Adjective + noun combination of nominal compound formation is almost not evident in the textbooks that the present study consulted.

5.2.1.1.4.2.2 Compound adjectives

In addition to the strategies of lexical elaboration via nominal compounding, Oromo is also found to enrich itself through creating new words by combining nouns and adjectives or noun and verb. The following data are collected from current Grade 8 biology and grade 11 Oromo textbooks.

(84) Base words	compound adjectives	gloss
a) <i>nama</i> 'man' + <i>tolche</i> 'made'→ <i>nam-tolchee</i>		'man-made'
<i>waaqa</i> 'God' + <i>uume</i> 'create'→ <i>waaq-uumee</i>		'God-created'
b) <i>amala</i> 'conduct' + <i>qabeessa</i> 'that possess'→ <i>amala-qabeessa</i>		'good conduct'
<i>kottee</i> 'leg' + <i>duudaa</i> 'unsplited hoof'→ <i>kotte-duudaa</i>		'equidae (odd-toed ungulate) mammal'

As seen in (84), compound adjectives in Oromo are formed from a noun and an adjective or a noun and as it also noted by Shimelis (2014:13). The compounds mentioned in (84a) are compound adjectives which are formed from noun and verb, where as the compounds in (84b) are composed from noun and adjectives as in *kottee* 'leg' + *duudaa* 'unsplited'→*kotte-duudaa* 'equidae (odd-toed ungulate) mammal'. In addition to the compound adjective formation that we have seen above, Temesgen (1993:40) claims that compound adjectives can also be formed via the combination of adjective and adjective as in *adii* 'white' + *qulqulluu* 'clean'→*adii qulqulluu* 'pure-white' and *ajaa'aa* 'stinky' + *tortoraa* 'rotten'→*aja'a tortoraa* 'rotten stinky'. However, such combinations are not compounds; but, adjective phrase. This is due to the fact that in compounds no elements can be inserted between modifier and head of a compound, while this is often possible for phrases (Ziering 2018:124). For instance, in *aja'a tortoraa* 'rotten stinky', we can insert word *baayyee* 'very' to make it *aja'a baayyee tortoraa* 'very rotten stinky'. In such compound, the modifier is not able to be modified as in, for instance, *sanbata guddaa* 'sunday'. To the contrary, in a syntactic phrase like *adii qulqulluu* 'pure-white', *qulqullu* 'pure' can be modified *bareedaa* 'beautiful' and become *adii qulqulluu bareedaa* 'beautiful pure-white'. Hence, this study argues that Oromo adjective compounds cannot be formed by the combination of adjective and adjective and such compounding strategy is not evident in the data sources this study consulted. Such combinations are possible in syntactic phrase.

5.2.1.1.4.2.3. Verbal compounds

Nominal and adjectival compounding are not the only lexical elaboration strategies of the language. Even though verbal compounding is not as productive as the nominal compounding, it is also one of the means by which Oromo enriches its lexicon. In, Oromo, verbal compounds are formed by the combination of ideophones and the verb *jedhe* 'say' and ideophones and another the verb *godhe* 'make'. From the perspective of semantics, ideophones represent actions like

breaking, entering, movement, hitting, eating etc. Ideophones also represent different sounds like human sound, animal sounds, sounds of different phenomena etc. Below are examples of Oromo verbal compounds which are formed via the combination of ideophones and the verb *jedhe* 'said'. The data are collected from grade 10 and 11 Oromo textbooks.

(85) Ideophone	gloss	verb	verbal compound	gloss
<i>bir</i>	'pass'	<i>jedh-Ø-eay-3MSG-PV</i>	<i>bir jedhe</i>	'pass suddenly'
<i>cal</i>	'quiet'	<i>jedh-Ø-esay-3MSG-PV</i>	<i>caljedhe</i>	'became silent'
<i>coloq</i>	'inter (water)'	<i>jedh-Ø-esay-3MSG-PV</i>	<i>coloqjedhe</i>	'fall into water'
<i>gad</i>	'down'	<i>jedh-Ø-esay-3MSG-PV</i>	<i>gadjedhe</i>	'tilt down'
<i>ol</i>	'up'	<i>jedh-Ø-esay-3MSG-PV</i>	<i>ol jedhe</i>	'stood up'
<i>qaa</i>	'snap'	<i>jedh-Ø-esay-3MSG-PV</i>	<i>qaa jedhe</i>	'snapping sound'

As indicated in (85), Oromo verbal compounds are orthographically open compounds (the constituent elements are spaced off, i.e., written as separate words) and not written as a single word or not even written hyphenated as in, for instance, the nominal compound *rog-sadee* 'triangle'. When verbal compounds in (85) are seen from the semantic point of view, all of them are endocentric as the meaning of the output emanated from the meaning of the inputs. Oromo verbal compounds do not allow any element to appear in between the ideophones and converbs. This is one of the criteria for compoundhood. In Oromo verbal compounds, it is the second constituent of the compound (the dummy verb) that accepts any inflectional or derivational morphemes. Oromo ideophones do not take any form of affix. This could possibly be due to the fact that most Oromo ideophones are of the imitative type (Shimelis 2014:70). The data in (86345) can clearly support the fact that word inflection undergoes on the head and not on the ideophone.

(86)a)	<i>Caalatuun</i>	<i>cal</i>	<i>jette</i>	
	<i>Caaltuu-n</i>	<i>cal</i>	<i>jedh-t-e</i>	[jette]
	Chaltu-NOM	IDEO	say-3SGF-PV	
	'Chaltuu keep quiet'			
b)	<i>Caalaan</i>	<i>cal</i>	<i>jechuu barbaade</i>	
	<i>Caalaa-n</i>	<i>cal</i>	<i>jedh-uu</i>	[jech-uu] <i>barbaad- Ø-e</i>
	Chala-NOM	quite	say-NMNZ	want- 3MSG- PV
	'chala wants to be quite'			

From the data in (86a), it is important to note that Oromo verbal compounds receive affixes at their heads (i.e., in the second constituent). Here, the second constituent of the verbal compound is inflected for gender and not the first constituent (i.e., the ideophone). In the same manner, the derivational affix *-uu* is attached in the same place (at the terminal position of the second constituent).

Similarly, the combination of Oromo ideophones and the verb *godhe* 'made' can result in verbal compounds. Such possible combinations are given in (87) below. The data are collected from grade 10 and 11 Oromo textbooks

(87) Ideophone	gloss + verb	→	verbal compound	gloss
<i>biliq</i> 'kill'	<i>godh-Ø-emake-3MSG-PV</i>		<i>bilq godhe</i>	'he killed'
<i>cim</i> 'silent'	<i>godh-Ø-emake-3MSG-PV</i>		<i>cim godhe</i>	'he made silent'
<i>girie</i> 'sip (flood)'	<i>godh-Ø-emake-3MSG-PV</i>		<i>girie godhe</i>	'he sipped'
<i>liqim</i> 'swallow'	<i>godh-Ø-emake-3MSG-PV</i>		<i>liqim godh</i>	'he swallow suddenly'
<i>ol</i> 'up'	<i>godh-Ø-emake-3MSG-PV</i>		<i>ol godhe</i>	'he vomited, stood up'
<i>qacam</i> 'caught'	<i>godh-Ø-emake-3MSG-PV</i>		<i>qacam godhe</i>	'he caught (suddenly)'

As can be seen from (87), most of the verbal compounds that are formed via the combination of ideophones and the verb *godhe* 'made' are semantically endocentric. The verbal compounds like *ol godhe* 'he vomited', however, exocentric as the meaning of the resultant compound do not emanated from its constituents. Just like the verbal compounds we have seen in (85), all compounds in (87) are open. The affixing takes place on the second constituent. The following

data shows the affixing of strategies of verbal compounds formed through the combination of ideophones and the verb *godhe* 'made'. The data are collected from grade 9 Oromo textbook.

(88)a)	<i>Badhaasaa ol godh-Ø-e</i>		
	Bedasa up	made-3MSG-Pv	
	'Bedasa vomitted'		
b)	<i>Badhaatun liqim gochuu barbaade</i>		
	<i>Badhaatu-n liqim godh-uu [goch-uu]</i>	<i>barbaad-Ø-e</i>	
	Bedasa-NOM swallow	made-NMNZ	want- 3MSG- Pv
	'Bedasa want to swallow food suddenly'		

As seen from (88), the verbal compounds are neither inflected nor derived at their first constituents; rather, it is the second constituents that take inflectional and derivational affixes as in (88a) and (88b) respectively. This is the feature of most of Oromo compounds.

5.2.1.1.4.3 Compounds of Oromo in view of the relations between their constituent elements

In relation to the classification of Oromo compounds that are found in the education sector, particularly in all textbooks of primary and secondary (only Oromo textbooks as subject) school, we have discussed the types of compounds based on the orthographic and categorical point of views in above subsections. In addition to the two criteria we have seen under the subtopic 5.2.1.1.4.1 and 5.2.1.1.4.2, the compounds of Oromo can also be classified into three: i.e. appositional, endocentric and exocentric. Such compounds of Oromo are discussed in the following subsection from the perspective of the relations between their constituent elements. The data in (89) are collected from grade 9 Oromo textbook.

(89) Base words	compounds	gloss
<i>barataa</i> 'student' + <i>barsiisaa</i> 'teacher' →	<i>barataa barsisaa</i>	'student and teacher'
<i>gurra</i> 'ear' + <i>laphee</i> 'heart' →	<i>gurraa laphee</i>	'ear and heart'
<i>harka</i> 'hand' + <i>miila</i> 'leg' →	<i>harkaa miila</i>	'hand and leg'
<i>qaama</i> 'body' + <i>qalbii</i> 'mind' →	<i>qaamaa-qalbii</i>	'body and mind'

In case of an appositional compounding, the semantic of the resultant compound is completely a compositional in which constituents are combined without the existence of interdependency (see also Shimelis 2014:202). In the appositional compounding, there is a terminal vowel length of the first constituent and this serves as conjunction *fi* 'and' as in, for instance, *qaamaa-qalbii* (lit. body and mind) 'due attention'. In such a compound, both constituents are heads (hence, such compounds are double headed compound).

There are also compounds in which one of the constituent being the head and the other being the modifier in Oromo as in *mana murtii* 'court'. In this example, the compound *mana murtii* (a house kind of house) is semantically the hyponym of its head called *man* 'house'. In endocentric compounds, its semantic emanate from its constituent. More illustrative examples of endocentric compounds are given in (89). There is also Oromo exocentric compounds whose meaning do not emanate from its constituents. Exocentric compounds do not have any semantic head as in, for instance, *abbaa caffee* 'chairman of the legislative assembly'. Several compound formations with the first constituent *abbaa* or *haadha* is also given by Tamene (2000:25, 134). The following data in (90) clearly show a number of Oromo exocentric compounds can be formed by combining *abbaa* 'father' or *haadha* 'mother' as a first element and another word as a second element. The data below are collected from grade 12 Oromo textbook.

(90) Base words	compounds	gloss
<i>abbaa</i> + <i>alangaa</i> 'lash'	→ <i>abbaa alangaa</i>	'prosecutor'
<i>abbaa</i> + <i>bokkuu</i> 'stick'	→ <i>abbaa bokkuu</i>	'traditional law-abiding person'
<i>abbaa</i> + <i>boora</i> 'white'	→ <i>abbaa booraa</i>	'owner of white horse'
<i>abbaa</i> + <i>caffee</i> 'grass'	→ <i>abbaa caffee</i>	'chairman of the legislative assembly'
<i>abbaa</i> + <i>duula</i> 'campaign'	→ <i>abbaa duulaa</i>	'traditional minister of war'
<i>abbaa</i> + <i>gadaa</i> 'year'	→ <i>abbaa gadaa</i>	'traditional president'
<i>abbaa</i> + <i>seera</i> 'law'	→ <i>abbaa seeraa</i>	'judge'
<i>abbaa</i> + <i>warraa</i> 'family'	→ <i>abbaawarraa</i>	'husband'
<i>haadha</i> + <i>warraa</i> 'family'	→ <i>abbaawarraa</i>	'wife'

There are a number of Oromo compounds that can be formed via the combination of *abbaa* 'father' and another constituent word. The same is true for the first constituent *haadha*; though not as productive as the previous constituent. The combination of *haadha* 'mother' and another word is not frequently used in the textbooks. As seen from (90), the semantics of such combination, most of the time, result in exocentric compounds. Such strategies of compound formation are not new methods that are evident in the textbooks that this study consulted. The strategy has been used in Oromo for a long period of time to designate traditional titles. In deed, the strategy is still being used to coin new words so as to enrich the language. For instance, the nominal compound *ija* 'eye' + *laafina* 'softness' → *ijala afina* 'fear' is purely exocentric compound whose meaning does not emanate from its constituent elements.

When we note the compounds in (90) from the perspective of their construction, they have a genitive nature of formations. The second constituents of all compounds stated under (90) have a genitive relation with the first constituent. Such genitive relations are supported with the terminal vowel length of the second constituents of the compounds. However, semantically, they serve as lexicalized compounds. This is due to the fact that the compounds represent a specific concept and additionally they fulfill the criterion such as lexical integrity, morphological and other criterion.

5.2.1.1.4.4 Complex compounding

In addition to the binary combination of words which we have seen in the previous subsections, Oromo enriched itself via complex compounding strategies. The combinations of three or more words are among the complex compounding strategies that are discussed under this subtopic. The following data which have the combination pattern of ((a + b) +c) are collected from grade 7 and 8 mathematics and biology textbooks respectively.

(91) Pattern of combination	compound	gloss
((a + b)+c)		
((<i>amala</i> + <i>jijjiirraa</i>) + <i>hammattuu</i>)	<i>amala jijjiirraa hammattuu</i>	'associative property'
'property change' 'bracket of'	(property + change + bracket of + change)	
((<i>amala</i> + <i>jijjiirraa</i>) + <i>iddoo</i>)	<i>amala jijjiirraa iddoo</i>	'commutative property'

'property change of'	'place'	(property + change + place of)	
((<i>amala</i> + <i>zeeroo</i>) + <i>ida'uu</i>)		<i>amala zero ida'uu</i>	'additive property of zero'
'property zero'	'addition of'	(property + zero + addition of)	
((<i>doonii</i> + <i>bishaan</i>) + <i>jalaa</i>)		<i>doonii bishaan jalaa</i>	'submarine'
'ship water'	'under of'	(ship + water + under of)	
((<i>qoricha</i> + <i>dhukkuba</i>) + <i>fursiisu</i>)		<i>qoricha dhukkuba fursiisu</i>	'aspirin'
'drug disease'	'analgesic of'	(drug + disease + analgesic of)	
((<i>seelii</i> + <i>dhiiga</i>) + <i>diimaa</i>)		<i>seelii dhiiga diimaa</i>	'white blood cell'
'cell blood'	'red of'	(cell + blood + red of)	

As can be seen from (91), the compounding strategy has to pass two processes so as to result in a complex compound. The first process is the formation of the first constituent of the complex compound from two words and the second process is the combination of the first constituent and the second constituent in order to form a complex compound as in, for instance, (*amala* 'property' *jijjiirraa* 'change of') + *hammattuu* 'bracket of')→((*amala jijjiirraa hammattuu*)). As the data in (91) shows, it is evident that Oromo textbook compilers used complex compounding strategy to designate scientific terminologies.

In Oromo, complex compounds can also be formed by the combination of three or more words; but, with a different construction (i.e., with the structure of (a+ (b +c))). The following illustrative examples of the form (a+ (b +c)) are collected from grade 7 and 8 chemistry and biology textbooks respectively.

(92) Pattern of combination	compound	gloss
(a + (b+c))		
(<i>dhaaba</i> + (<i>bu'uura</i> + <i>misoomaa</i>))	<i>dhaaba bu'uura misoomaa</i>	
(organization (base+development))	(organization+base+development)	'organization of infrastructure'
(<i>makaa</i> + (<i>gosa</i> + <i>addee</i>))	<i>maka agos-addee</i>	
'solution of' 'type multiple'	(solution+ type +multiple)	'solution of multiple type'
(<i>makaa</i> + (<i>gosa</i> + <i>tokkee</i>))	<i>makaa gos-tokkee</i>	
'solution of' 'type multiple'	(solution + type + single)	'solution of single type'
(<i>sibiila</i> + (<i>alkaalaayini</i> + <i>dachee</i>))	<i>sibila alkaalayini dachee</i>	
'iron + alkaline earth of'	(iron + alkaline+ earth)	'alkaline iron of earth'
(<i>sirna</i> + (<i>bullaa'insa</i> + <i>soorataa</i>))	<i>sirna bulla'insa soorataa</i>	

'system' 'digestion food of' (<i>sirna</i> + (<i>marsaa</i> + <i>dhiigaa</i>))	(system + digestion + food of) <i>sirna marsaa dhigaa</i>	'digestion system'
'system' 'blood circulation of' (<i>xaa'oo</i> + (<i>nama</i> + <i>tolchee</i>))	(system'+ 'blood+ circulation of) <i>xaa'oo nam-tolchee</i>	'blood circulation system'
fertilize' 'man' 'made'	(fertilizer + man- made)	'man-made fertilizer'

Looking into the data in (92), one may conclude that Oromo can have a complex compound formation. Complex compounding strategy is also used during textbook preparation as a means of enriching the lexical gaps of the language. There are a number of scientific and technical terms which are expressed via complex compounds and such means are evident in the textbooks this study consulted. Even though the number of combining words is the same with the number of combining elements of the data in (91), the complex compounds in (92) are different in the structure they are formed. For instance, the complex compound *sirna bullaa'insa soorataa* 'digestion system of food' or simply 'digestion system' has to pass a two-step combination process. In the first process, the second constituent of the complex compound *bullaa'insa soorataa* 'digestion of food' (a two word compound) has to be formed by the combination of the noun *bullaa'insa* 'digestion' and another noun *sirna* 'system'. The second process is to combine the noun *soorataa* 'food' with the already formed nominal compound *bullaa'insa soorataa* 'digestion of food' so as to result in the complex compound *sirna bullaa'insa soorataa* 'digestion system'.

The complex compounds stated in (91) and (92) can also be noted from their elements' modifier and modify as well as from specifier and specify relationships. In (91), *bishaan* 'water' is the specifier of noun *doonii* 'ship' in the binary compound *doonii bishaan* 'ship of water'. In the complex compound *doonii bishaan jalaa* 'submarine', *jalaa* 'under of' is specifier of the first constituent (i.e. the binary compound *doonii bishaan* 'ship of water') of the complex compound. Whereas in (92), the entire binary compounds (the second constituents of the respective complex compounds) are the modifier of the first constituents (the words) of the complex compounds. For instance, in the binary compound *bullaa'insa soorataa* 'digestion of food' (which itself is the second constituent of the complex compound *sirna bullaa'insa soorataa* 'digestion system'), *soorataa* 'food of' is the modifier of the second constituent *bullaa'insa* 'digestion'. Further more,

bullaa'insa soorataa 'digestion of food' is still modifier of the first constituent (noun) *sirna* 'system' in the complex compound *sirna bullaa'insa soorataa* 'digestion system'.

Apart from the formation of compounds via combining two or three words, there are also several Oromo complex compounds which are coined from four words. There are two compounds of binary words within such complex compounds. The following illustrative examples in (93) clearly show that Oromo has a rich possibility of coining a number of complex compounds of four constituent in order to describe scientific concept. The complex compounds stated under (93) are collected from grade seven and eight chemistry as well as grade eight mathematics textbooks.

(93) Pattern of combination	compound	gloss
$((a + b) + (c + d))$		
$((ashaboo + atoomii) + (gosa + lamee))$	<i>ashaboo atoom-gos-lamee</i>	'salt of two type atom'
'salt' 'atom' 'type' 'two of'	(salt of +atom-type-having two)	
$((boba'oo+ haftee)+ (lubbuu + qabeeyyii))$	<i>boba'oo haftee lubbu-qabeeyyii</i>	'fos fuel'
'fuel' 'fossil' 'life' 'that possessil	(fuel+ fosil + life-that possess)	
$((roga + sadee) + (kofa sirrii))$	<i>rog-sadee kofa sirrii</i>	'right angle triangle'
'side' 'having three' 'angle' 'right'	(side having three + angle right)	

The mathematical and scientific terms are described using complex compounds of four constituents as in *rog-sadee kofa sirrii* 'right angle triangle' and *caasaa gabatee peeredikii ammayyaa* 'structure of modern periodic table' respectively. The complex compounds in (93) can be represented using combination structure $((a + b) + (c + d))$. Oromo complex compounds could also be formed by combining loanword constituents with the endogenous words of the language as in, for instance, *kompaawundii atoomii gos-lamee* 'a two type compound atom'; a complex compound formed from four constituents.

There are even more complex compounds which are evident in the textbooks that this study consulted. The following complex compounds which are formed through the combination of five or more words are collected from grade eight biology textbook.

(94) Pattern of combination and the resultant complex compounds.

a) ((*orgaanizimii* + (*seelii* + *qeenxee*)) + (*boca* + (*qophe* + *siliipparii*)))

'organism of' 'cell' 'single' 'shape' 'shoe' 'slipper'

orgaanizimii seel-qeenxee boca qophe siliipparii

'slipper shaped unicell organism'

b) ((*sirna* + *wal-hormaata*) + ((*saala* + *malee*) + *baakteeriyaa*))

'system' 'reproduction' 'sex' '-non (without)' 'bacteria of'

sirna wal-hormaata saal-malee baakteeriyaa

'non sexual reproduction system of bacteria'

The data stated in (94) are given as an illustrative example to show how textbook compilers used very complex compounding strategies so as to describe scientific terminologies. In such complex compounding method there are three or more compounds within the given complex compounds. For instance, in (94a), we can have a binary compounds such *seel-qeenxee* 'unicell', *qophe siliipparii* 'slipper'. We can also compounds that have three constituents as in *orgaanizimii seelii qeenxee* 'unicellular organism' and *boca qophe siliipparii* 'shape of slipper'. The complex compounding methods applied in (94a) and (94b) can be represented by structures ((a+ (b+c)) + (d+ (e+f))) and ((a+b) + ((d+e) +f)) respectively.

Complex compounds of Oromo could also be noted from their orthography perspective. As we have seen the orthography of binary constituent of simple compounds in section 5.2.1.1.4.2, the compounds formed from two constituents could be classified in closed, open and hyphenated compounds as in, for instance, *caalbaasii* 'auction', *jecha diigala* 'compound word' and *miil-soboo* 'false feet' respectively. Unlike a two constituent Oromo compounds, all complex compounds of the language are open compounds (i.e., this study could not get complex compounds of Oromo which are written as a single word being hyphenated or closed). To the contrary, there are complex compounds which are written as a single word (closed) in Hadiyya as in ((*leho* + *hoora*) + (*heechcha* + *hoora*)) → *lehoorheechchoora* (death-deprive-life-deprive) 'very difficult situation' as noted by Samuel (2017:168). To him, some compounds which were originally open compounds have been changed into closed compounds through transitional changes. There are also closed Oromo compounds which are formed by combining two elements

(words) as in *caasa* 'structure' + *luga* 'language' → *caasluga* 'grammar'. However, in Oromo, it is difficult to conclude that language's closed compounds might have been open and became closed as the time passes. This study recommends a further diachronic investigation to arrive at the truth.

Looking into the elements (constituents) of Oromo complex compounds in (91), (92) and (93), one can also note that some complex compounds of the language are combined from Oromo indigenous and borrowed words so as to describe foreign concept. For instance, the complex compound *orgaanizimii seel-qeenxee boca qophees siliipparii* 'slipper shaped unicell organism' is formed through the combination of the indigenous words *qeenxee* 'single', *boca* 'shape' and *qophees* 'shoe' as well as the borrowed words *orgaanizimii* 'organism' *seel* 'cell' *siliipparii* 'slipper'. The defining features of Oromo compounds that we have seen under this chapter have the same features with those that have been described in chapter four (see section 4.1.2.3).

5.2.1.1.5 Blending

Blending is a term that refers to a linguistic unit which is composed of elements that function independently in other circumstances (Crystal 1991). It is the merger of two words into one to mean a new concept. Ahn (2014:1) noted blending as a word-formation process in which two or more independent words are merged into a new word with the shortening of at least one of the source words. Blending is a combination of the parts of two words, usually the beginning of the word and the end of another. There are, however, exceptions to this as it can be seen in the following sections. The data in (95) are collected from the current grade 9-12 Oromo textbook and from Qajeelcha 1998 compiled by the standardization committee of Oromo which is being used as a reference in secondary schools. The numbers written in parentheses in front of the base words and blends refer to the number of syllables of respective base words and blends.

(95) Base words	blending	gloss
(a) <i>dura</i> (2) 'front' + <i>galaa</i> (2) 'that inter'	→ <i>durga</i> (2)	'preposition'
<i>duuba</i> (2) 'back' + <i>galaa</i> (2) 'that inter'	→ <i>duubga</i> (2)	'postposition'
(b) <i>gamisa</i> (3) 'semi' + <i>duubbachiisaa</i> (4) 'vowel'	→ <i>gamduu</i> (2)	'semi vowel'

As shown in (95a), there are blends in Oromo which are formed through clipping without overlapping. To make the analysis of the issue simpler, we can represent the parts of the first word as WX and parts of the second word as YZ. All blends in (95a) and (95b) are formed via the combination of the first part of the first word and the first part of the second word (i.e. WX + YZ→WY). When the two base words do not have a shared segment, the syllable structure plays a role in determining the switch point. The blends in (95a and 95b) are coined through the combination of the first syllable of the first base word and from the onset and the nucleus of the first syllable of the second base word. For instance, in (95b), the blend is formed via the combination of the first syllable of the first word and through the omission of the coda of the first syllable of the second word (constituent) i.e., *gamisa* 'semi' + *duub.bach.'ii.saa* 'vowel'→*gamduu* 'semi vowel'. From the data in (95), it can be noted that the first base words have contributed more segments than the second base words so as maximize the semantic transparency of the blend. There are, however, some Oromo new blends which are formed contrary to the combination in (95b). The data in (96) which are collected from Qajeelcha (1998) clearly show the fact.

(96)Base words	blend	gloss
<i>latii</i> (2) 'morpheme' + <i>wabala</i> (3) 'bound'→	<i>lawab</i> (2)	'bound morpheme'
<i>latii</i> (2) 'morpheme' + <i>walaba</i> (3) 'free' →	<i>lawal</i> (2)	'free morpheme'

As shown in (96), there are some instances in which blends are formed through the combination of the onset and the nucleus of the first base word and the first syllable (that onset, nucleus and coda) of the second base word as in *latii* 'morpheme' + *wabala* 'bound'→*lawab* 'bound morpheme'. Here, in data (96), unlike data in (95) in which the first base word contributed more segments, the second base words contributed more segments than the first base words to maximize the semantic transparency of the blends. There are, however, instances in which both base words contribute equal segments as in *lama* (2) 'two' + *hidhii*(2) 'lip'→*ladhii* (2)'bilabial'. The following data are collected from the current grade 11 and 12 Oromo textbook as well as from Qajeelcha (1998).

(97) Base words	blend	gloss
<i>a) antee</i> (2) 'relative' + <i>hima</i> (2) 'sentence'		→ <i>antima</i> (3)'object'
<i>caasaa</i> (2) 'structure' + <i>hima</i> (2) 'sentence'		→ <i>caasima</i> (3)'syntax'
<i>gocha</i> (2) 'action' + <i>hima</i> (2) 'sentence'		→ <i>gochima</i> (3) 'verb'
<i>leexaa</i> (2) 'single' + <i>hima</i> (2) 'sentence'		→ <i>leexima</i> (3) 'simple sentence'
<i>mataa</i> (2) 'head' + <i>hima</i> (2) 'sentence'		→ <i>matima</i> (3)'subject of a sentence'
<i>b) tajaajila</i> (4) 'service' + <i>hima</i> (2) 'sentence'		→ <i>tajima</i> (3) 'case'
<i>xaxaa</i> (2) 'complex' + <i>hima</i> (2) 'sentence'		→ <i>xaxima</i> (3)'complex sentence'

The data in (97) show that the word *hima* 'sentence' is often found to come as the second constituent irrespective of any word it combines with to form a blend. As presented in (97), the left source words of blends are longer than (regarding their segments or sounds) their corresponding right source words. Here, it can be noted from the data in (97) that since the right source word determines the length of the whole blend, more segments could be preserved from each source word. This is due to the fact that, as noted by Bat-El 2006:66, there may be a suspect that the formation of blends is not accidental, but rather governed by some general principles. The principles reflect two competing tendencies: (i) to truncate segments from the base in order to allow the blend to have the length of a single word, preferably one of the base words, and (ii) to preserve as many segments from the base words as possible and thus maximize the semantic transparency of the blend. When the data in (97) are noted from the perspective of the numbers of syllables, all blends contain more syllables as compared to their respective base words with the only exception of the case of (97b) *tajaajila* 'service' (a quadri-syllable base word) i.e. the base words have two syllables whereas the blends contain three syllables. There are also other Oromo blends which are formed (but not many) by the truncation and combinations of three base words. Still, there are blends which are the results of truncation of only one of the base words taking all segments of the other base word. The data in (98) which are collected from the current grade 12 Oromo textbook and from a reference book called Qajeelcha (1998) can help to show the fact mentioned above.

(98)	Base words	blend	gloss
a)	<i>dabala</i> (3) 'add' + <i>gocha</i> (2) 'action' + <i>hima</i> (2)'sentence	→ <i>dabalgochima</i> (5)	'adverb'
b)	<i>eeyyee</i> (2) 'yes' + <i>miti</i> (2) 'no' →	<i>eemiti</i> (3)	'paradox'
	<i>jecha</i> (2) 'word' + <i>maka</i> (2) 'mixed' →	<i>jemaka</i> (3)	'blend'
	<i>latii</i> (2) 'morpheme' + <i>yaasa</i> (2) 'derivation' →	<i>layyaasa</i> (3)	'derivational morpheme'
	<i>mataa</i> (2) 'head' + <i>hima</i> (2) 'sentence' →	<i>matima</i> (3)	'subject'
	<i>xinxala</i> (3) 'study' + <i>gooqa</i> (2) 'language' →	<i>xinqooqa</i> (3)	'linguistics'
	<i>xinxala</i> (3) 'study' + <i>madda</i> (2) 'source' →	<i>xinmadda</i> (3)	'anthropology'
c)	<i>xinxala</i> (3) 'study' + <i>yaada</i> (2) 'idea' →	<i>yaaxxina</i> (3)	'theory'

The blend formation process in (98) has something which makes it different from the rest of all blend formation. The blend formation process that we have seen in data (95) follow WX+ YZ→WY (96) and (97) follow WX+ YZ→WZ, most of the formation in (98b) WX + YZ→WYZ whereas blends in (98c) are formed contrary backward combination i.e. WX + YZ→ YW. The data in (98a) clearly prohibit us from concluding that all blending in Oromo is formed via the combination of only two base words as in, for instance, the blend *dabalgochima* 'adverb' is coined by truncating three base words i.e. *dabala* 'add', *gocha* 'action' and *hima* 'sentence'. Although they are not many, Oromo blends are also being seen to be formed by truncating the segments of only the first base word and taking all segments of the second base word as in *jecha* 'word' + *maka* 'mixed' → *jemaka* 'blend'. On the other hand, Bat-El (2006:66) noted that blends in which only the first word undergoes truncation could also be considered a clipped compound, especially when each word contributes only one syllable to the surface form, which is a characteristic of clipped compounds.

Blends in Oromo can also be formed through the overlapping of segments of the base words. The following data collected from grade 11 Oromo textbook clearly shows the above fact.

(99)Base words	blend	gloss
<i>gocha</i> (2) 'action' + <i>haala</i> (2) 'situation'	→ <i>gochaala</i> (3)	'mood'
<i>rawwii</i> (2) 'complite' + <i>hima</i> (2) 'sentence'	→ <i>raawwima</i> (3)	'present perfect tense'
<i>hidhii</i> (2) 'lip' + <i>ilkee</i> 'dental'	→ <i>hidhilkee</i> (3)	'labiodental'

As seen in (99), some Oromo blends are formed through overlapping of element(s) of the base words. For instance, *gochaala* 'mood' is formed via preserving *gocha* from *gocha* and *haala* from *haala*. Here, we can see that in *gochaala* 'mood' all segments of the base word appear in the blend. But, there are exceptions in which segments are truncated from the base words in existence of overlapping of segments of the base words as in *rawwii* 'complite' + *hima* 'sentence'→*raawwima* 'present perfect tense'. On the other hand, it is true that the more segments of the base words in the blend there are, the more transparent the base words are. The selection of the position of the shared segment(s) as the switch point contributes to segmental maximization. The shared segments overlap and thus correspond to segments in both base words, allowing more segments from each word to be preserved in the blend. We can also note that the presence of overlapping segments in a blend can have an effect of maximizing the number of segments of its source words, as can be seen in (99). The bold segments *a* and *i* in (99) are segments from both source words.

With regard to the number of syllables in a blend, it can be noted that, in most of the blends that we have seen under the present subtopic, the number of syllables in blends is greater than the number of syllables in base words with the exception of the blends in data (95), (96),and in (97b). When measured with the number of syllables, most of the blends are longer than their base words. In (95), the size of the syllables in the blends is equal to the number of syllables in all base words except the case of 95b in which the number of syllables in blends is less than the base words. In (96), the size of the syllables in the first base word and blends are equal.

Blends in Oromo can also be noted from the perspective of the switch point of the blend (i.e. where the first source word ends and the second begin). Accordingly, in the absence of a shared segment, the switch point is determined by the syllable structure. The switch point of blends of Oromo do not follow similar patterns in any size of the syllables of a blend (bi-syllabic, tri-

syllabic or polysyllabic blends). For instance, in disyllabic blends of Oromo like *duubga*, *lawab* and *ladhii*, the switch points are at coda–onset, nucleus-onset and nucleus-coda boundaries respectively (i.e. *duub.a* 'back' + *gal.aa* 'that inter' → *duub.ga* 'postposition', *lat.ii* 'morpheme' + *wab.ala* 'bound' → *law.ab* 'bound morpheme' and *lam.a* 'two' + *hidh.ii* 'lip' → *ladh.ii* 'bilabial' respectively). The switch point of trisyllabic blends like *antima*, *eemiti* and *xinqooqa* (*an.tee* 'relative' + *him.a* 'sentence' → *an.tim.a* 'object', *eey.yee* 'yes' + *mit.i* 'no' → *eem.it.i* 'paradox', and *xin.xala* 'study' + *qooq.a* 'language' → *xin.qooq.a* 'linguistics') respectively are at onset-nucleus, nucleus-onset and at syllable boundaries respectively. On the other hand, when there is an overlap, the switch point is determined by the shared segments, which overlap in the blend as in *gocha* 'action' + *haala* 'situation' → *gochaala* 'mood' (see data 99).

Regarding semantic relations of base words and the resultant blends, as can be seen from all data that we have in this subtopic, Oromo blends deliver the concept of their two (in some the cases three) base words and their meanings are thus contingent on the semantic relation between the base words. Most of Oromo blends are endocentric as all base words play a role on the semantics of the blend or their meaning results from the combination of the meaning of the base words. To put it in another way, most Oromo blends are made up of base words where they denote something which has the feature of both elements. Their semantic feature will make them similar with compounds of Oromo as there are endocentric compounds in which both constituents contribute to the meaning. Since Oromo blends are seen to maximize its size (as it can be noted from the data under this subtopic that the majority of them obtain more syllables than the base words), they tend to facilitate the semantic recoverability of the base words. It can also be noted that most of the blends that are discussed under this subtopic are blends that are related to language area terminology.

Oromo blend formation processes can also be seen from the perspective of the faithfulness of input to output and output to output correspondence. As noted by Ahn (2014) and Bat-el (1996), the input-output correspondence requires faithfulness of the output to its input. Conversely, the output-to-output correspondence requirement is the faithfulness between output forms. As it can be seen in most of the blending process of Oromo in present subtopic, truncated forms and its base form are faithful to each i.e., there is an output to output correspondence. For instance,

during the blend formation of the case of *gocha* 'action' + *hima* 'sentence' → *gochima* 'verb', the outputs (truncated forms) are faithful among themselves (i.e., the underlined truncated forms (outputs) are similar to the blend (output)). Oromo base words and the resultant blends are in surface-to-surface correspondence. Regarding blend formation in Oromo there is considerable debate in literature when only one of the constituent truncated and attached to the second word as authors claim such formation could result in blend or compound. But, there is no debate of classifying the resultant formation in to compound or blend when there is lose of materials from all constituents since such formation blending. Tamene (2000:12), however, classifies the formation of new words via truncation of segments of both constituents as in, for instance, he gives the formation *agarsiisa* + *dhugoo* > *agargoo* 'data' to be compound and not blend. This study argues such formations are not compounding; but blending.

5.2.1.1.6 Acronyms and Abbreviations

Abbreviation involves shorting of the existing words to create other words, usually the informal versions of the original words. Oromo textbook compilers used the abbreviation method as mean of covering the lexical gaps. The following data collected from grade nine & ten Oromo as well as grade eight biology and mathematics textbooks.

(100) Full form	abbreviation	gloss
a) <i>yookin</i>	ykn	'or'
<i>fakkeenya</i>	Fkn	'e.g.'
b) <i>bilbila</i>	bil.	'Tel.'
<i>keeyyata</i>	keey.	'paragraph'
<i>lakkoofsa</i>	lakk.	'number'
c) <i>tartiiba lakkoofsaa</i>	T.L.	'serial number (S/N)
<i>lakkoofsa bilbila</i>	L.bil.	'Tel.'
c) <i>kilomeetira</i>	km	'km'
<i>digirii seelshiyeesii</i>	°C	°C
<i>digirii faaranaayitii</i>	°F	°F
<i>meetir iskuweerii</i>	m2	m2

As can be seen in (100), Oromo new words are formed in the textbooks via by shortening the existing words. Such kind of new word formation is clipping as seen in (100a). A single abbreviated word can be created through incorporating non-initial letter of the existing word. For instance, the new word *ykn* is coined through eliminating the internal parts of the existing word *yokkin* 'or' and the same is true for the formation of *Fkn* 'example'. The abbreviation method employed in (100b) is different from that of (100a). All the three new words in (100b) are formed by clipping (eliminating) the last parts of the respective words as in *keey* from the word *keeyyata* 'paragraph'. The new word formed by process of lopping off a portion and reducing it to a monosyllabic or disyllabic rump has the same meaning as the original lexical terms. Here, we have to note that the fragment does not have to be the salient part of the original word, like the case of (100b), neither prosodically nor semantically.

New words are also coined in textbooks by the process of initialism from a two word compound as in *T.L* from *tartiiba lakkoofsaa* 'serial number (S/N)'. An acronym is the process whereby a new word is formed from the initial letters of the constituent words (i.e. via recursive initialism). We can also note that Oromo can abbreviate and form a new word through initialism and clipping as in *lakkofsa bilbla* → *L.bil* respectively (i.e. from initial word *lakkoofsa* the initial letter *L* is taken and the second word *bilbila* 'phone' is clipped as *bil*). Oromo textbook compilers have also borrowed the word and the compounds together with their respective abbreviations as seen in (100d).

In addition to the abbreviations and acronyms formed from words and simple compounds, Oromo is found to enrich itself by initialism (alphabetism) which is an abbreviation pronounced by reciting the individual letters as in the English initialism cases of *British Broadcasting Corporation-BBC*, *Digital Video Disc -DVD*, *Portable Document Format- PDF*. The following abbreviations and acronyms are collected from grade eight mathematics; grade 10 and 11 Oromo textbooks.

(101) Full form	acronyms	gloss
<i>akka lakkoofsa habashaa</i>	A.L.H.	'E.C (Ethiopian Calendar)'
<i>akka lakkoofsa Itoophiyaa</i>	A.L.I	'E.C (Ethiopian Calendar)'
<i>akka lakkoofsa Awuroppaa</i>	A.L.A	'G.C (Gregorian Calendar)'
<i>hirmaataa walii xiqqichaan</i>	H.W.X	'LCM (Least Common Multiple)'
<i>kan kana fakkaatan</i>	kkf	'et cetera'
<i>lakkoofsa sanduqa postaa</i>	L.S.P	'PO box'
<i>waan kana fakkaatu</i>	wkf	'et cetera'

Textbook compilers have also abbreviated from three words. Oromo abbreviations that are detected in the textbooks are written in different ways. As can be seen in (101), some acronyms are written with a capital letter and a full stop/period in place of the deleted part to show the ellipsis of letters as in *akka lakkoofsa habashaa*→A.L.H to mean 'E.C (Ethiopian calendar)'. In some other cases, however, words are abbreviated with small letter and with no full stop/period in between the letters as in *waan kana fakkaatan*→wkf and *kan kana fakkaatan* kkf both to mean 'et cetera'.

Acronyms are also formed from complex compounds and such abbreviations are found in the textbooks that this study consulted. The following data are collected from grade eight mathematics, grade five, seven, and nine as well as grade eleven Oromo textbooks.

(102) Full form	acronyms	gloss
<i>akka lakkoofsa Awroppaa</i>	ALAW	'G.C (Gregorian Calendar)'
<i>biiroo barnoota Oromiyaa</i>	BBO	'OEB (Oromia Education bureau)'
<i>dhaabbata dimokraatawaa ummata Oromoo</i>	DhDUO	'Oromo people democratic organization'
<i>dhaabbata fayyaa addunyaa</i>	DhFA	'WHO (World Health Organization)'
<i>dhaloota Kiristoos booda</i>	Dh.K.B	'AD (Anno Domini) after the birth of Christ'
<i>dhaloota kiristoos dura</i>	Dh.K.D	'BC'
<i>fooyyessa qulqullina barnoota waliigalaa</i>	FQBW	'general quality improvement of education'

<i>gama harka bitaa</i>	GHB	'left hand direction'
<i>hirmaataa walii guddicha</i>	HWG	'GCM 'greater common multiple)'
<i>intarpiraayizoota daldaala xixiqqaa</i>	IDX	'micro trade interprise'
<i>yuunivarsiitii madda walaabuu</i>	Y.M.W	'MWU) 'Mede Welabuu Unicersity'

There are also a number of new words of Oromo which are formed from the initial letters of the constituent words as in (102). Except the third example in (102) (i.e., DhDUO 'Oromo people democratic Organization'), all acronyms are pronounced as sequences of letters (as an alphabet). The acronym *DhDUO* ('Oromo People Democratic Organization') will, however, not be pronounced letter by letter; but via applying regular reading rules of words as [dhadi'o] in a similar case of pronouncing the English acronym OPEC as [oupek] (See also section 4.1.3). There are serious orthographic variations of Oromo abbreviations in the textbooks. Such variations are discussed in the orthographic variation sections of this chapter.

Apart from the lexical elaboration of the language via new word formation by abbreviating indigenous words, there is also abbreviation of loanwords that are incorporated into Oromo, particularly in the textbooks. The following data are collected from grade mathematics as well as grade seven and eight social study textbooks.

(103) Full form	Acronymization/abbreviation	gloss
<i>Addi Bilisa Baasaa Ummata Ertiraa</i>	EPLF	'Eritrea People Libration Front'
<i>Addi Bilisa Baassaa Ummata Tigraay</i>	TPLF	'Tigray People Libration Front'
<i>Baankii Misoomaa Afrikaa</i>	ADB	'African Development Bank'
<i>Dhaabbata Nyaataafi Qonna Addunyaa</i>	FAO	'Food and Agriculture Organization'
<i>Kiloomeetira</i>	km	'kilometer'
<i>Komiishiinii Dingadee Afrikaa</i>	ECA	'Economic'
<i>PaartiiDimookraataawaaOromoo</i>	ODP	'Oromo Democratic party'
<i>Seentimeetira</i>	cm	'centimeter'
<i>Doktar</i>	Dr.	'doctor'

Looking in into the data in (103), one can note that, in addition to the lexical loanwords, Oromo has incorporated several abbreviated loanwords into its lexicon so as to enrich itself. Here, we can note that all abbreviated loanwords are used in Oromo exactly as they are being in English. Except the two, *km* and *cm*, all are formed via taking the initial letters of each word in the source language (English). On the other hand, Oromo has also included the clipped (words created through incorporating non-initial letter) loanwords of English as in, for instance, *km* and *cm*. Oromo, for instance, did not use *OLF* (Oromo Liberation Front) i.e. the English version of the Oromo abbreviation *ABO* (*Adda Bilisummaa Oromo*) in various texts that this study consulted. Instead, indigenous abbreviated Oromo word, *ABO*, is used despite the existence of its English version.

Rather than incorporating and using abbreviated and clipped loanwords, as shown in (103), Oromo could have better elaborated its lexicon via internal method by abbreviating and clipping all the examples stated above with the exception of the abbreviation *km and Dr*. Each of the above example could be represented as *ABBUE*, *ABBUT*, *BMA*, *DNQA*, *KDA*, *PDO* and *sm* respectively. Hence, it can be concluded that, there are cases in Oromo in which abbreviated loanwords are taken being in similarly as they are abbreviated in the source language; but being written in full form in Oromo as seen in (103) under the first and the second column. Abbreviation in Oromo is also seen to be formed in the textbooks via taking the initial and the final letters of a word, as in, for instance, *Dr* <*doktar*> 'doctor'. The above abbreviated loanwords are pronounced similarly as they are pronounced in the source language (English) except that of *km*, *cm* and *Dr*. Such abbreviations are read in their full form as seen in the first column in (103). On the other hand, there are also abbreviated loanwords in Oromo being pronounced as a word, as in, for instance, *AIDS* <*Eedsiin*>.

5.2.1.2 Orthographic variation of Oromo in the textbooks

The main task of education, concerning language related issues, is to present and promote "the standard language" (Rubin 1973:2). The use of a standard language in education helps to avoid differences of skill and knowledge to be acquired at various education levels. The very critical reason for the use of standard language in education is for the sake of effective communication and simplicity.

Going back to the above sub topic, due to the absence of any guiding rule, words, compound words, abbreviations etc. are being spelled in various ways. It seems, now, customary to spell separate words sometimes together and in some other cases with spaces and hyphen in between them. There are also cases in the textbook in which Oromo words are not spelled as it they are not read. For instance, we (the mother tongue speaker of the language) read the English words 4% as *parsantii afur* 'percent four' not as *afur parsantii* 'four percent. With the same case, the way it is spelled and how it is read in Oromo is different the way it is spelled and read in Amharic and English (*arat pərsənt* and four percent respectively). In Oromo, however, 4% is read (spelled) as *parsantii afur* 'percent four' but not as *afur parsantii* 'four percent' as opposed to the case of Amharic and English. Hence, such discrepancy of orthography and the way word or words are spelled and other issues of variations in the language are discussed in the sections to come via taking practical examples from the textbooks.

There are no overtly known conventions as to how Oromo compounds should be written. This could be the main reason for the inconsistencies of writing Oromo compounds. The standardization committee of Oromo seems to forget such critical issue of standardizing the language. The following data show inconsistently written Oromo compounds and the respective rule based propositions. The data are collected from grade six and grade eight mathematics, grade seven and eight biology and chemistry as well as from grade 9-12 Oromo textbooks.

(104) Inconsistently written form	proposed form	rule of proposition	gloss
a) <i>atoombaay'ee</i> / <i>atoom-baay'ee</i> /	<i>atom baay'ee</i>	deletion of the terminal vv of 1 st word	
	<i>atoom baay'ee</i>		'multiatom'
b) <i>ba-maqaa/bamaqaa</i>	<i>bamaqaa</i>	deletion of the ccv of 1 st word	'pronoun'
c) <i>giddu-galeessa/giddu galeessagiddu galeessa</i>		deletion of the terminal vv of 1 st word	'average'
d) <i>keessa-deebii/keessadeebii/</i>	<i>keessa deebii</i>	leaving the two words independent	'revision'
	<i>keessa deebii</i>		
e) <i>lubbu-qabyyee/lubbu qabyyee/</i>	<i>lubbuqabeeyyii</i>	deletion of the terminal v of 1 st word	'living'
	<i>lubbuqabeeyyii</i>		
f) <i>miil-soboo /miil soboo</i>	<i>miilsoboo</i>	deletion of the terminal v of 1 st word	'false leg'
g) <i>nam- tolchee /namtolchee</i>	<i>namtolchee</i>	deletion of the terminal v of 1 st word	'man-made'

-
- h) *qub-guddeessaa/ qubguddeessa* deletion of the terminal vv of 1st word 'capital letter'
qubguddeessaa
- i) *rog-sadee/rog sadee/rogsadee rogsadee* deletion of the terminal vv of 1st word 'triangle'
- j) *seelqeenxee/seel-qeenxee/ seelqeenxee* deletion of the terminal vv of 1st word 'unicell'
seel qeenxee
- k) *tarmi lamee /tarmi-lamee/ tarmilamee* deletion of the terminal v of 1st word 'biterm'
tarmilamee
-

As can be seen from (104), various Oromo compounds are inconsistently written as there is no convention as to how such compounds shall be written. The same compounds are orthographically spelled differently as hyphenated, with and without space in between the constituent compounds. For instance, the Oromo compound *rogsadee* 'triangle' is spelled in three different ways as *rog-sadee/rog sadee/rogsadee* in various textbooks and even within a single textbook itself. The proposition given by this study is, however, based on rules of compounding. The proposed orthography for the differently spelled compound *rog-sadee/rog sadee/rogsadee* 'triangle' is *rogsadee* 'triangle'. This is due to the fact that the compound is formed by the combination of the word *roga* 'side' and another word *sadee* 'having three'. If there are deletions of segments of one of the constituents (mostly deletion of few terminal segments of the first constituent) of the compound, it is recommended to spell the compound as a single word as in *rogsadee* 'triangle'. Spelling the closed compound *rogsadee* 'triangle' as an open compound *rogsadee* or as a hyphenated compound *rog-sadee* do not make writers orthographically efficiency at least for two reasons. Firstly, the main reason for deleting the terminal segment (the vowel) *a* of the first constituent *roga* is to spell the compound as a single word. Secondly, adding either white space or hyphen via deleting the terminal vowel of the first constituent creates an extra burden for writers.

Compounds of Oromo are also seen to be spelled in various ways in the absence of deletion of any segments of the constituents of compounds. If there is no elision of any segment of the constituents during compound formation, the constituents of the compound could better be spelled independently with white space in between them. For instance, the compound *keessa deebii* 'revision' is spelled in different ways as in *keessa-deebii/keessadeebii* and *keessa deebii*. Such compounds could better be spelled as any open compounds (with white space in between

the constituents) like *Abbaa Duulaa* 'traditional minister of war' since the constituents of such compounds are independent.

Hence, here, one can note that, except the compound *keessa deebii* 'revision', all the inconsistent forms of spelling of compound words occurred due to the deletion of segments of the first constituent of the compound and unnecessary and not economical addition of either a white space or a hyphen in between the constituents. One can also note that compounds of Oromo are written inconsistently due to the lack of conventions as to how they shall be written. There seems also either loose or no communication between the textbook compilers and the committee. Therefore, the standardization committee of Oromo has to codify (set rules or conventions) so as to establish a standard method of writing compounds. Textbook compilers and editors have to be aware of language area conventions. There should also a strong link between the committee and textbook compiler so as to minimize such inconsistencies.

There are also variations while writing abbreviation in the textbook corpus that this study consulted. The main aim of abbreviation is to save space, comply (to obey to usages and orders) with the already set rules by language authoritative bodies and to avoid repetition. Abbreviations in Oromo are being used inconsistently throughout all the textbooks which this study consulted. The following data are collected from grade eight mathematics and from grade five, seven, nine as well as from grade eleven Oromo textbook to show the existing abbreviation usage variations within and among textbooks.

(105) Full form	variation in initialization/abbreviation	gloss
<i>Akka lakkoofsa Awroppaa</i>	ALAW/ALA/A.L.A	'G.C (Gregorian Calendar)'
<i>Akka lakkoofsa habashaa</i>	A.L.H./ALH	'E.C (Ethiopian Calendar)'
<i>Akka Lakkoofsa Itoophiyaa</i>	A.L.I/ALI	'E.C (Ethiopian Calendar)'
<i>Akka lakkoofsa Awuroppaa</i>	A.L.A/ALA	'G.C (Gregorian Calendar)'
<i>Biiroo barnoota Oromiyaa</i>	B.B.O/BBO	'OEB (Oromia Education bureau)'
<i>Dhaabbata dimokraatawaa ummata Oromoo</i>	DhDUO/Dh.D.U.O	'Oromo People democratic Organization (OPDO)'
<i>Hirmaataa walii xiqqichaan</i>	H.W.X/HWX	'LCM (Least Common Multiple)'
<i>Kan kana fakkaatan</i>	k.k.f./ kkf	'et cetera'
<i>Keeyyata</i>	keeyy./ keey.	'paragraph'
<i>Lakkoofsa</i>	lak./Lak./ /lakk./Lakk.	'number'
<i>Lakkoofsa saanduqa poostaa</i>	LSP/L.S.P	'po.box'
<i>Okkoofsa bilbila</i>	lakk.bilbilaa /L.bil.	'Tel.'

Apart from the orthographic variations that have been discussed under the compound, there are also orthographic variations of abbreviations. As seen in (105), various abbreviations are being spelled in different ways as there are set rules by any concerned body. The existence of guideline as to how and which kind of word or words to be abbreviated or not, plays a great role in minimizing the orthographic variation of abbreviations in a language. What we observe from data in (105), however, is different as a single abbreviation is, now being spelled in two or more ways. For instance, *lakkoofsa* 'number' has been spelled in four various ways in the textbooks as *lak./Lak./ /lakk./Lakk.* through the incorporation of non-initial letters. Here, the orthographic variation arises due to usage of lower and upper cases and size of the incorporated letters. In some other cases, as in *B.B.O/BBO*, for instance, the variation arises as a result of the period in between the initial letters of the abbreviation. The standardization committee of Oromo should prepare a guideline (manual) so as to facilitate the usage of one standard spelling for a particular abbreviation. Otherwise, it is clear that such orthographic variation affects the standardization process of the language and the learners as well as users of the language. The standardization committee of also has to set rules that show when to abbreviate (if there are good justification to abbreviate) and not to abbreviate.

There are also orthographic variations in writing abbreviations of measurements. The following data are collected from grade nine and eleven Oromo as well as from grade eight biology and chemistry textbooks.

(106) Full form	abbreviation	the way it is read	gloss
<i>dhibba digirii seelshiyesii</i>	100 °c	<i>digirii seelshiyesii dhibba</i>	'hundred degree celsius'
<i>jaatama degree</i>	60°	<i>digirii jaatama</i>	'sixty degree'
<i>kudhan digirii faaranaayitii</i>	10°F	<i>digirii faaranaayitii kudhan</i>	'ten degree fahrenheit'
<i>kudhan sentimetirii/a</i>	10 cm	<i>sentimeetrii kudhan</i>	'ten centimetre 10cm'
<i>sadi parsantii/parsantii sadi</i>	%3/ (3%)	<i>parsantii sadi</i>	'three percent (3%)'
<i>sagal kilomeetira</i>	9 km	<i>kilomeetira sagal</i>	'nine kilometre (9km)'

As seen in (106), there are variations while using the measurement such as percentages, kilometer etc. For instance, % 'percent' is being spelled in various ways; sometimes before and in other cases after numbers as in 3% and %3 in the textbooks. Oromo adapted abbreviated loan measurement units and incorporated them into its lexicon without any adjustment. For example the adapted loanword, *cm* could better be abbreviated in Oromo as *sm* < *senti meetira* > 'centimeter' if and only if the loan word *senti meetira* 'centimeter' is to be used despite the existence of the genuine Oromo word *dhibbeentaa* for the same concept. It is obvious that words have to be read in the order they appear so as to minimize variation and to regularize orthography and pronunciation in the language. In this regard, one can also note here that the abbreviated loanwords with numerical values stated under the second column of (106) will not be read in exactly the same way as they are written. For instance, the orthography of *9km* directs us to be read as <*sagal kilomeetrii*> 'nine kilometre'. This is, however unacceptable, in Oromo. The acceptable way of reading *9km* is as <*kilomeetrii/ra sagal*>. Here, the mismatch between the orthography under the second column and the way they are read under the third column in (106) is due to the non-adjustment of abbreviated loans while borrowing from other languages i.e. English. Hence, to make the orthography and the way they are spelled consistently, abbreviated loanwords (measurement terms (signs)) should come before numerical values as in, for instance, %3 <*parsantii sadi*> 'three percent', *km 9* <*kilomeetrii/a sagal*> 'nine kilometre' etc.

There is also a failure of textbook compilers to use the codified words (expressions) that are codified by the standardization committee of Oromo. For instance, the word *dhibbeenta* is codified by the authority to designate the concept 'percent'. But textbook compilers used the form *parsantii* to represent the same concept. The word *parsantii* 'percent, %' should not be used as far as the genuine word of the language *harka* (*dhibbeentaa*) exists. This could be an evidence for the loose relationship between the standardization committee of Oromo and textbook compilers as it has been reported by the committee itself during focus group discussion session.

In addition to the orthographic variation of abbreviations in Oromo discussed above, there are also other orthographic variations at the word level. The following illustrative data are taken from grade three, six, nine and eleven Oromo textbooks.

(107) Variation	gloss
<i>ammoo/ immoo</i>	'however'
<i>ishee/ishii/isii</i>	'she'
<i>televiziyoonii/ televiziina/televiijinii</i>	'television'
<i>xinnoo/ xiqqoo</i>	'little'
<i>yookiin/yookaan</i>	'or'

There are several instances in which a single Oromo concept represented by a number of forms (spellings). As can be seen from data in (107), the same concept in Oromo is being spelled in various ways in the textbooks due to the lack of codification by the standardization committee of Oromo. In such representation of a single concept with more than two spellings as in, for instance, *ishee/ishii/isii* can negatively affect the standardization process of the language. Such variation, except for the last example, arises due to dialect variation of the language. The use of two or more alternative spellings for the same word in the textbook will obviously affect learning to read the language and learning as a whole particularly at the lower grades

On the other hand, in some few cases, as in the case of loan word *televiziyoonii/ televiziina/televiijinii* 'television', orthographic (spelling) variation may occur due to variation in representation of a single sound with more than two graphemes. Here, the sound /z/ in the loan word 'television' is represented in the textbooks with various graphemes; <zn>, <zh>, <z>, <zy>

and <j>. This variation in representation (graphization) itself leads to difficulties for the learner during phonics instruction which helps learners familiarize themselves with the orthography of Oromo. It is evident that there are several instances in which various forms are used interchangeably to represent a single concept throughout the textbooks consulted. It is obvious that the usage of standard and firmly fixed forms in every textbook and in all education levels avoids learners' misunderstandings and uncertainties. Hence, the standardization committee of the language to codify only one of the existing forms so as to standardize the lexicon of the language.

There are also other orthographic variations apart from the spelling variation that we have seen at the word level. The following illustrative examples taken from grade six, nine and eleven Oromo textbooks in order to show the existing variations that resulted from the attachment and detachment of a word from another word.

(108) Inconsistently written form	gloss
<i>addaafi eenyummaa/ addaa fi eenyummaa</i>	
culture.and identity / culture and identity	'culture and identity'
<i>ofirraa darbachuu / of irraa darbachuu</i>	
self.on throw/ self on throw	'throwing away (from oneself)'
<i>kun rajeeffanno hintanedha / kun rajeeffanno hin taane dha</i>	
this exclamation NEG.become.COP/this exclamation NEG become COP	'this is not exclamation'

Looking into the data in (108), one can note that there is variation while writing words. Although the standardization committee of Oromo codified the conjunction marker *fi* and the copula *dha* to be spelled together with the preceding, they are being inconsistently spelled in the textbooks (Qajeelcha 2008:1-2). They are sometimes written together with the preceding word and in some other places it is seen to be spelled independently in the textbooks as in, for instance, *addaafi eenyummaa/ addaa fi eenyummaa* (culture.and identity/ culture and identity) 'culture and identity' and *kun immoo ittisudha/ kun ammo ittisu dha* (this however protection. COP/ this however protection COP) 'this is, however, protection'. Here, the conjunction *fi* 'and' and copula *dha* spelled in the first case together with the preceding word. To the contrary, in the second case, they are spelled independently. In the same document of the standardization committee or

Oromo, the negation marker *hin* is codified to be spelled independently without attaching it to the word that come after it.

There are also a number of variations in Oromo primary and secondary textbooks as to how two and more words are spelled. In this regard, the standardization committee of Oromo did not still state a clear rule to write some words which are being spelled in different ways as seen in (108). For instance, two independent Oromo words are spelled in two ways as in *ofirra/of irra* 'on oneself' and *kun raajeeffanno hin-ta'-n-edha/ kun raajeeffanno hin ta'-n-e dha* (this exclamation NEG-become-NEG-PVCOP/this exclamation NEG become-NEG-PV COP) 'this is not exclamation'. Here, two words are spelled as a single word and as two independent words as in *of irra/ofirra*'on oneself'. There are also cases in Oromo in which three words written as a single word and as three independent as in *hin-ta'-n-e-dha/hin ta'-n-e dha* (NEG-.become-NEG-PV-COP/ NEG become-NEG-PV COP) respectively.

Since there are gaps of conventions as to how such division of words shall be spelt, except for the conventions stated, all various possible forms are free to be used or be written based on the preference of language users. If textbook compilers, the authors other individuals and institutions used any form they like, then there will be misunderstanding among users of the language. The standardization committee of Oromo, therefore, has to continue setting rules and developing a guideline so as to minimize such variation in the textbooks. This is due to the fact that the use of standard language in education (in the textbook in the context of this study) helps to deliver a standard form of knowledge at various levels of education. Hence, Oromo textbook compilers have to adhere to the rules, spelling conventions and guidelines set by the committee (though there are several issues still to be regularized) in order to maintain the standard language starting from early lower grades. Here, it is obvious that textbooks are the most important educational materials to rich the massive parts of a society. Unless all the concerned bodies play their own role to minimize variation during the preparation of textbooks, it is difficult to regularize the orthography, lexicon and other aspects of the language.

There are also other variations in the textbooks that are related to lexicons. Challenges of language standardization could also come from the existence several varieties within the language. Due to very slow processes of standardization of the language and loose

communication between standardization committee of the language and the education sector expertise in general and textbook compilers in particular, the process of norm selection and usage seems challenging as one variety is used at one time and some other variety at another time. Such cases are clearly evident in the textbook corpus that this study consulted. The data in (109) are collected from the current grade four and five environmental and social study textbooks respectively as well as from grade five, nine, ten and eleven Oromo textbooks to show the lexical variations in textbooks that needs normalization. The examples of lexical variation are given below are based on Feda's (2015) objective classification of varieties of Oromo.

(109)Western	Northern	Central	Southern	Eastern	Southeastern	Gloss
<i>arfaasaa</i>	<i>furma</i>	<i>afraasaa</i>	<i>badheessa</i>	<i>badheessa</i>	<i>arfaasa</i>	'spring'
<i>daraara</i>	<i>daraaraa</i>	<i>abaaboo</i>	<i>daraaraa</i>	<i>ilillii</i>	<i>daraaraa</i>	'flower'
<i>kijiba</i>	<i>kijiba</i>	<i>soba</i>	<i>kijiba</i>	<i>kijiba</i>	<i>dhara</i>	'false'
<i>nafa (dhanna)</i>	<i>jismii</i>	<i>dhagna</i>	<i>qaama</i>	<i>nafa</i>	<i>nafa</i>	'body'
<i>waanjo</i>	<i>jeedallo/jajalloo</i>	<i>geedala</i>	<i>ukkullo/jeedal'o</i>	<i>ukkullo</i>	<i>sardiida</i>	'fox'
<i>woggaa</i>	<i>umrii</i>	<i>ogгаа</i>	<i>ganna/umrii</i>	<i>ammata</i>	<i>ganna</i>	'age'

As can be seen in (109), there are inconsistent ways of using words in the textbooks stated above. Though the standardization committee of Oromo opted the composite approach of normalizing (regularizing) the language, the committee did not codify such lexical variations via selecting only one word as a norm and stating other forms of the same concept as its variety. The textbook compilers did not consistently apply everywhere in the textbooks only one form to a single concept. Rather, they applied one form at one time and other forms in some other time. All words in (109) are not consistently used in the textbooks consulted. For instance, four forms are evident to be used for the same concept 'spring' as in *arfaasaa*, *furma*, *afraasaa* and *badheessa*. There is no doubt that the use of such words *inconsistently* throughout the textbooks will definitely affect one of the main goals of language standardization i.e. minimal variation in form.

On the other hand, this study does not mean that Oromo has to have only one form (spelling) and one pronunciation for each lexicon as this is the ideal code. But, what has to be emphasized, here, is that any form of a language (be its phonology, morphology, lexicon or grammar) has to possess (exhibit) high degree of stability (regularity) to which various writers (i.e., textbook

compilers, authors and journalists) refer to so as to minimize misunderstanding. Going back the examples given under (109), though some dialect of the language use the same form for the given concept as in, for instance, *nafa* in Western, Eastern and southeastern dialect for the concept ‘body’, Oromo has several lexical forms for a single concept which is being used interchangeably in the textbooks as in the last case in same example.

Let alone the lexicons that are not yet standardized (normalized) by the standardization committee of Oromo, there are evidences in the textbooks that can clearly show that what are being codified are not being put into practice by the education sector (in our case by the textbook compilers). The data which are taken from *Wiirtuu 1995 volume 1* and grade four Oromo as well as from grade three environmental science in (110) can depict the destandardization tendency of the language in the textbooks

(110) SCO	application in textbooks	gloss
<i>Dilbata</i>	<i>Dilbata, Sanbataguddaa</i>	'Sunday'
<i>Jimaata</i>	<i>Jimaata, Juma'a</i>	'Friday'
<i>Kamisa</i>	<i>Kamisa</i>	'Thursday'
<i>Kibit'ata</i>	<i>Kibxata, Lammaffo</i>	'Tuesday'
<i>Roobii</i>	<i>Roobii, Yaato</i>	'Wednesday'
<i>Sanbata</i>	<i>Sanbata duraa, Sanbata, Sanbata xinnaa</i>	'Saturday'
<i>Wiixata</i>	<i>Wixata, Dafinoo, Hojjaduree</i>	'Monday'

The inclusion of words from various varieties of a language could be regarded as strategy of accommodation of varieties. In such a means, standardization process follows the composite approach of regularizing a language. This approach can increase the acceptance of the standardization process and the norm by the speaker of the language. The application of the selected forms from various varieties has to be, however, consistent throughout written materials. Strong application of codified forms of phonology, lexicon, morphology and grammatical units of a language in the textbooks, on the other hand, could obviously play a paramount role in the standardization process of a language.

What can be seen in (110), however, is against the principles of standardization as which has been regularized is not being put into practice. Here, one can note that textbook compilers working against the standardization committee of Oromo as they used their own words to represent a concept, despite the existence of codified words by the authority. For instance, as seen in (110) under the first column, the authority clearly codified the names for the days of a week. Except for *Thursday* in which *Kamisa* is uniformly used throughout the textbook corpus, textbook compilers have not consistently applied the codified words for the names of the days of the weeks as seen in (110) under the second column. When we see the inconsistent application of more than three forms for a single concept in textbooks as in, for instance, *Wixata*, *Dafinoo* *Hojja duree* for 'Monday' and *Sanbata duraa*, *Sanbata* and *Sanbata Xinnaa* for 'Saturday' and so on, it is not difficult to imagine how such challenges negatively affect the race for standardization of the language.

The challenges of implementation of codified words in the key area of enforcement (in the education sector) are not limited to what have been discussed so far. There several words which are normalized by the authority, but denied to be applied in the textbooks. The four main seasons of the year are clearly codified; but they are not consistently applied in the textbook corpus. For instance, *Spring* is codified as *Birraa* but the words *Arfaasaa* and *Badheessa* are used to refer to the same concept 'Spring' (grade 5 Social study textbook 2013:29). On the same document and page, *Birraa* is used not to refer to *Spring*; but for the concept *Autumn*.

The preparation of learning materials (textbooks) obviously requires a strict and consistent application of codified language so as to play a positive role in the race of standardization of a language. Still, there are other evidences from textbooks which prove the existence of inconsistent use of the codified (normalized) words. For instance, the names of months of the year are among the concepts that are codified by the authority (Standardization committee of Oromo). But; as the data in (111) indicate, textbook writers inconsistently applied in the textbook corpus. The data are collected from the standardization committee of Oromo *Wiirtuu* volume1; grade 5 and 7 social study and grade 11 Oromo textbooks.

(111) SCO	application in textbooks	gloss
<i>Adoolessa</i>	<i>Adoolessa</i>	'July'
<i>Ammajii</i>	<i>Ammajii, Furma</i>	'January'
<i>Bitootessa</i>	<i>Bitootessa</i>	'March'
<i>Caamsaa</i>	<i>Caamsaa, Maazaa</i>	'May'
<i>Ebla</i>	<i>Ebla, Badheessa</i>	'April'
<i>Fulbaana</i>	<i>Fulbaana, Birra</i>	'September'
<i>Gurraandhala</i>	<i>Gurraandhala, Badheessa</i>	'February'
<i>Hagayya</i>	<i>Hagayya</i>	'August'
<i>Muddee</i>	<i>Muddee, Arfaasaa</i>	'December'
<i>Onkoololessa</i>	<i>Onkoololessa</i>	'October'
<i>Sadaasa</i>	<i>Sadaasa, Birraa</i>	'November'
<i>Waxabajjii</i>	<i>Waxabajjii</i>	'June'

The data in (111) clearly shows that the majority of the codified words for the names of months of the year are not consistently applied in the textbook. In some of the cases, textbook writers seem to mix up the names of the months with the names of the season of a year, as in, for instance *Sadaasa* and *Birraa* for *November* as well as *Mudde* and *Arfaasaa* for *December*. Textbook writers seem also to use a word from their own dialects to refer to the names of a month as in *Caamsaa* and *Maazaa* for *May*. From the above data, some of the codified names of months of the year are consistently used throughout the textbooks as in, for instance, *Onkoololessa* 'October', *Bitootessa* 'March' and the rest. The challenges of implementation of the codified forms obviously hinder the standardization process of the language as the use of standard language in the education sector, particularly in the various learning materials (textbooks) is a key site to promote the codified and standard form. It is true that minimum variation in form (spelling) can exist in learning materials. Education sector, however, has to use the form that avoids miscommunication and promote regularized language. This sector has an important impact to fix the codified form and it can also destandardize the language.

There are also other several inconsistent applications of forms in the textbooks. The following data show the destandardization tendencies as what are codified by the standardization

committee of Oromo are not put into practice. The data in (112) are taken from grade ten and twelve Oromo, grade three and four environmental science, grade seven mathematics as well as from grade seven biology textbooks.

(112) SCO	application in textbooks	gloss
<i>dhamjecha</i>	<i>dhamjecha, latii</i>	'morpheme'
<i>hanqaaquu</i>	<i>hanqaaquu, buuphaa, killee</i>	'egg'
<i>lukkuu</i>	<i>hindaaqqoo, lukkuu</i>	'hen'
<i>ogummaa</i>	<i>ogummaa, dandeettii</i>	'skill'
<i>qarshii</i>	<i>qarshii, birrii</i>	'birr'
<i>sadakkuu</i>	<i>rogsadee</i>	'triangle'
<i>wallee</i>	<i>muuziqaa</i>	'music'

As can be noted from (112), there are misapplications of even technical terms like *dhamjecha* 'morpheme' and *sadakkuu* 'triangle' that are codified the legitimate authority (the standardization committee of Oromo). Here, we can note that the codified and the uncoded terms are being used interchangeably by textbook writers. There are also more interesting things in (112). The textbook writers used the uncoded forms through rejecting the codified terms and this is an extreme case which seems a task of destandardizing the language.

All in all, from the document (textbook corpus) analysis, we can note that the lexical standardization process is identified to be challenged by the inconsistent application of the codified (normalized) forms in the textbooks. A concept in the textbooks is represented in Oromo in a word (form) from one variety at one time and in a word (form) from another variety at another time. There are several inconsistently applied words in the textbook corpus that the present study consulted. The level of inconsistency starts from using various alternative forms for the same concept to rejecting what has been codified by the authority. Unless there exists a strong link between the authority and the education sector in general and textbook writers in particular, the standardization process is almost impossible as educational materials are the major resources via which the learning society accesses to the standard and normalized language. It is also through textbooks that the standard language is transmitted to the society at large. To the

contrary, if a single concept in the language is expressed using various forms (spelling), as we have seen in the above consecutive examples, is nearly impossible to see a consistent application of words (forms) in the current textbooks at all. This in itself again creates high variation in form which leads to misunderstanding. Hence, if such action is continued, then the language will continue to be destandardized.

5.2.1.3 External means of lexical elaboration

Apart from the internal means of lexical, external methods of lexical enrichment can also be used after fully applying the lexical items of that language. External means of elaborating a language, however, is used if and only if there is no internal means of covering the lexical gaps of a language. Accordingly, borrowing and loan translation as a means of external lexical elaboration of Oromo in the textbook corpus are presented and discussed in the sections to come.

5.2.1.3.1 Borrowing in textbooks

Borrowing is another source which textbook writers used as a means of lexical elaboration. Borrowing is the process of importing linguistic items from one linguistic system into another, a process that occurs any time two cultures are in contact over a period of time (Hoffer 2002:1). Loanwords are most of the time the result of cultural contact (intercultural communication). It is also obvious that the main function (among other functions) of a loanword is communicating the new object/action/idea (Hoffer 2002:1). The author also noted that in most of the cases loanword enters the language and becomes part of the regular dictionary. In some other cases, a culture may find (opt) an indigenous designation from its own for a new concept that entered into its culture.

5.2.1.3.1.1 Loanwords and loanword adaptation in the textbook corpus

As per the data collected from the textbooks, Oromo has borrowed from several endogenous and exogenous languages so as to enrich its lexicon during textbook preparation. All the loanwords have undergone a set of modification in order to conform to phonological systems of Oromo as we shall see. The source languages and the extent borrowing are discussed in following section.

5.2.1.3.1.1 Amharic loanwords in the textbook corpus

There are several Amharic words that are included in the lexicon of Oromo so as to cover the lexical gap. Amharic loanwords are introduced into Oromo during textbook preparation. The data in (113) show some of Amharic loanwords that are included in the lexicon of the language by the textbook writers. The data are collected from grade five social science study, grade eight chemistry and civics, grade one environmental science, grade seven and eleven Oromo as well as from grade six mathematics textbooks.

(113) Amharic		Oromo	Gloss
<i>ቤተክርስቲያን</i> /betəkɪrɪstɪyan/	→	<i>bataskaana</i>	'church'
<i>ቦይ</i> /boy/	→	<i>boyii</i>	'canal'
<i>ብርቱካን</i> /bɪrɪtkʷuan/	→	<i>burtukaana</i>	'orange'
<i>ደቦ</i> /dabbo/	→	<i>daabboo</i>	'bread'
<i>ደብተር</i> /dəbtər/	→	<i>dabtara</i>	'exercisebook'
<i>ደንብ</i> /dənb/	→	<i>danbii</i>	'regulation'
<i>ደቂቃ</i> /dək'ik'a/	→	<i>dakiqaa</i>	'minute'
<i>ደንች</i> /dɪnʃ/	→	<i>dinnicha</i>	'potato'
<i>ግብር</i> /gɪbɪr/	→	<i>gibira</i>	'tax'
<i>እብቅ</i> /ʔibq/	→	<i>habaqii/ibbiqii</i>	'dirty materialinteff'
<i>ቀረጥ</i> /k'ərət'/	→	<i>qaraxa</i>	'excise, duty'
<i>ቁምጣ</i> /k'umt'a/	→	<i>qumxaa</i>	'shorts'
<i>ሰንበት</i> /sənbət/	→	<i>sanbata</i>	'Sunday'
<i>ወቅት</i> /wək'it/	→	<i>waqtii</i>	'season'

As can be seen from (113), there are several Amharic loanwords which are introduced into the lexicon of Oromo during textbook preparation. Here, we can note that all the loanwords in (113) are incorporated into Oromo during textbook preparation with modification. Due the phonological difference between Amharic and Oromo, loanwords are adjusted so as to suit to

phonological features of the recipient language (i.e. Oromo). For instance, the non-existing phoneme Amharic *i* has to be changed (modified) in either $i \rightarrow u$ or $i \rightarrow i$ and $i \rightarrow a$ as in ብርካት */birtik^wuan/* 'orange' to *burtukaana*, ግብር *gibir* 'tax' to */gibira/* and ደብተር */dabtər/* 'exercise book' to *dabtara* respectively. There are also other modification which includes the addition of *i* and *a* at the end of the loanword as in ቦይ */boy/* 'canal' to *boyii* and ሰንበት */sənbət/* 'Sunday' to *sanbata*. There are also other adaptation mechanisms to entertain loanwords in the language. There are cases in which Oromo changes the terminal short vowel of the loanwords of source languages into long vowels as in, for instance, ደቂቃ */dək'ik'a/* in to *daqiiqaa* 'minute'.

5.2.1.3.1.1.2 English loanwords in the textbook corpus

Oromo has borrowed several words from English as a result of modern civilization and technological advancement. English is the major source of lexical elaboration for Oromo as compared to other endogenous and exogenous languages. This fact has been attested in the textbook corpus consulted in this study. Loanwords of English in Oromo are of several types; scientific, mathematics, biology, chemistry. Some of such loanwords of English and the phonological adaptation are discussed in the following subsections. Here, there may be two ways by which English loanwords to enter into the lexicon of Oromo. Some English loanwords could have integrated into Oromo via an intermediary language (Amharic) as Amharic has been language of administrative, medium of instruction, media etc. before probably any of Ethiopian languages. On the other hand, educated Oromo bilingual could play their own role in integration of some of English loanwords as the textbook developers are expected to be educated.

The data in (114) presents loanwords of English which are introduced into the lexicon of Oromo during textbook preparation. The following data are taken from grade one environmental science textbooks.

(114) English/gloss	Oromo	vowel replacement, insertion, addition
electric /ilektrɪk/	→ <i>elektirikii</i>	/i/→/e/, /ii/ addition
stove /stouv/	→ <i>istoovii</i>	/ov/→/oo/, /ii/ addition
machine /məʃɪn/	→ <i>maashinii</i>	/ə/→/aa/, /ii/ addition
mobile /moʊbi:l/	→ <i>mobaayiliii</i>	/ov/ →/o/, /i:/→/aa/, /ii/ addition/
radio /reɪdiʊ/	→ <i>yaadiyoo</i>	/ei/→/aa/, /iʊv/→/iyoo/
television /telɪvɪʒn/	→ <i>televiziinii</i>	/i/→/e/, /ii/ addition
tractor /træktər/	→ <i>tiraktera</i>	/æ/→/a/, /ə/→/e/, /a/ addition

As can be seen from (114), there are several electronic related terms that are borrowed from English and introduced into the lexicon of Oromo. Most of the loanwords stated in (105) are the name of electronic materials that are used in a house. All loanwords in the above data are introduced into Oromo via some phonological modification. Hence, there are phonological adjustments in all of the above loanwords. For instance, there are non-existent English vowels in the loanwords that undergo modifications to conform to the phonological pattern of Oromo as in *maashinii* 'machine' /məʃɪn/ where the non-existing vowel /ə/ is replaced by the existing. There are also other phonological adaptations Oromo in (114); vowel length /aa/, /æ/→/a/, /ə/→/e/ and /i:/→/aa/. Here, it can be noted that a single vowel and diphthong of English is seen to be replaced by long vowel of Oromo as in machine /məʃɪn/ by *maashinii* (/ə/→/aa/) and radio /reɪdiʊ/ by *raadiyoo* (/ei/→/aa/).

There are also English loanwords in Oromo which are related to mathematical terminology. There are several modifications of the syllable pattern while introducing the loanwords of English into Oromo. The data in (115) are collected from grade three mathematics textbook.

115) English/gloss	Oromo	vowel replacement, insertion, addition
digit /didʒit/	→ <i>dijitii</i>	/ii/ addition
diameter /daiæmitər/	→ <i>diyaameetirii</i>	/aiæ/→ /iyaa/, /i/→/ee/, /ə/→/i/ & +/ii/
fractions /frækʃns/	→ <i>firaakshinoota</i>	/æ /→/aa/, /i/ insertion
square /skwer/	→ <i>iskuweerii</i>	/e/ →/ee/ & cluster modification
lattice /lætis/	→ <i>kaatesii</i>	/æ/→/aa/, /i/→/e/ + /ii/ addition
the compass /ðə/ /kʌmpəs/	→ <i>koompaasicha</i>	/ʌ/→/oo/, /ə/→/aa/
perpendicular/pɜ:rpændikjələr/	→ <i>parpendikulaarii</i>	/ɜ:/→/a/, /ə/→/e/, /ii/ addition
problems /prɑ:bləms/	→ <i>pirobileemota</i>	/ɑ:/→/o/, /ə/→/ee/, /i/ insertion
proportionality/prəpɔ:ɾʃənæləti/	→ <i>piropporshinaalumm aa</i>	/ə/→/o/, /ɔ:/→/o/, /æ/→/aa/, /i/ insertion
radius /reidiəs/	→ <i>raadiyesii</i>	/ei/→/aa/, /iə/→/iyee/ + /ii/
rectangle /rektæŋgl/	→ <i>reeктаangilii</i>	/e/→/ee/, /æ/→/aa/, /ŋ/→/n/, /i/ insertion + /ii/
rhombus /rɑ:mbəs/	→ <i>rohombosii</i>	/ɑ:/→/o/, /ə/→/o/ , /o/ insertion + /ii/
trapezium /trəpi:ziəm/	→ <i>tiraappiiziyemii</i>	/ə/→/aa/, /iə/→/e

It is true that Oromo borrowed a number of mathematical terms from exogenous language (English) so as to cover its lexical gaps. All the above loanwords passed at least one phonological adjustment. There is the addition of /ii/ at the end of almost all loanwords as in *dijitii* 'digit', *iskuweerii* 'square' and so on. The replacement of English diphthongs with the Oromo long vowel, as in /ei/→/aa/ is also one of the phonological adaptation seen in (115). The non-existing English vowels are also seen to be replaced by the existing Oromo vowels to conform to the phonology of Oromo as in, for instance, /ə/→/aa/, /ə/→/ee/, /ɜ:/→/a/, /æ/→/aa/, /ʌ/→/oo/ and /ə/→/oo/. Apart from the addition and replacement of vowels, there are also English diphthongs Which are adapted through the insertion of glide y while they are introduced into Oromo as in *raadiyesii* radius /reidiəs/ /iə/→/iyee/, *tiraappiiziyemii* 'trapezium'/trəpi:ziəm/ /iə/→/iye/ and *diyaameetirii* diameter /daiæmitər/ /aiæ/→/iyaa/. Another major phonological

process seen under the above data is the adaptation of consonant cluster. Oromo allows two consonant clusters medial; but not initially and never more than two anywhere. This consonant cluster is, however, permissible in English. Hence, the impermissible consonant clusters in the loanwords are adapted via inserting *i* where there are impermissible consonant clusters as in, for instance, *firaakshinoota* fractions /frækʃns/ and *reeктаangilii* rectangle /rektæŋgl/ word medial and initial respectively.

On the other hand, looking into some of the above data, one can note that Oromo has most likely has a tendency of replacing its own inflectional and derivational morphemes as in *firaakshinoota* 'fractions' and *koompaasicha* 'the compass' (inflectional morpheme of number and definite marker respectively). Derivational morphemes of source language (English) are also seen to be replaced by the recipient language (Oromo) as in, for instance, proportionality /prəpɔ:ʃjənæləti/ → *piropporshinaalummaa*. Here, *-ummaa* is an Oromo noun forming derivational morphemes like in *nama* 'human', *nama* + *-ummaa* → *namummaa* 'humanity' which replaces the English derivational morpheme /əti/.

Oromo borrowed terms related to measurement. The following data presents some of mathematical or measurement related loanwords that are introduced into the language during textbook preparation. These data are taken from grade three mathematics textbook.

(116)English	Oromo	adaptation (replacement/insertion/addition)
gram /græm/	· <i>giraama</i>	/æ/ → /aa/, /i/ insertion, /a/ addition
kilogram /kiləgræm/	· <i>kiiloogiraama</i>	/ə/ → /oo/, /æ/ → /aa/, /a/ addition
metre /mi:tər/	· <i>meetira</i>	/i:/ → /ee/, /ə/ → /i/, /a/ addition
milimetre /milimi:metər/	· <i>miliimeetira</i>	/ɪ/ → /i/, /ə/ → /i/, /a/ addition

As we can see from the above data, several measurements related lexical items have been integrated into Oromo during textbook preparation by the textbook compilers. Although the main intention of presenting the data under (116) is to present measurement related technical terms, we have also several phonological adaptation that attracted our attention. Loanwords are adapted to the recipient's phonological features via replacement, insertion and addition processes as in,

for instance, *giraama* 'gram' /græm/ /æ/→/a/, /i/ and /a/ respectively. It can be noted that all the non-existing English vowels in (116) are seen to be replaced by the existing vowels of Oromo.

Apart from the inclusion of electronic, mathematical, measurement, Oromo is attested to enrich itself via borrowing words that are directly related to biological terminologies so as to express concepts. The following loanwords are collected from grade seven biology textbooks to show some of biological terms introduced into the language by the textbook writers.

(117)English	Oromo	adaptation process (replacement/insertion/addition)
aspirin /æsprɪn/	→ <i>aspiriini</i>	/æ/→/a/, /ɪ/→/ii/, /i/ addition
iron /æɪərn/	→ <i>ayireenii</i>	/æɪə/→/ayi/, /ee/, /ii/ addition
disinfectant /dɪsɪnfektənt/	→ <i>disinfeektaantii</i>	/ɪ/→/i/, /e/→/ee/, /a/→/aa/, /ii/ addition
scurvy /skɜːrvi/	→ <i>iskaarvii</i>	/ɜː/→/aa/, /i/ insertion, /ii/ addition
carbohydrate /kɑːrboʊhaɪdreɪt/	→ <i>kaarboohayidiretii</i>	/ɑː/→/aa/, /oʊ/→/oo/, /aɪ/→/ayi/, /eɪ/→/e/
organisms /ɔːrgənɪzəms/	→ <i>orgaanizimoota</i>	/ɔː/→/o/, /ə/→/aa/, /ɪ/→/i/, /ə/→/i/
ovary /oʊvəri/	→ <i>ovaarii</i>	/oʊ/→/o/, /ə/→/aa/, /i/→/ii/
protein /prəʊtiːn/	→ <i>pirootiinii</i>	/oʊ/→/oo/, /i/ insertion, /ii/ addition
cells /sels/	→ <i>seeloota</i>	/e/→/ee/,
vitamin /vɪtəmiːn/	→ <i>vaayitaaminii</i>	/aɪ/→/aayi/, /ə/→/aa/, /ɪ/→/i/, + /ii/

As can be seen in data (117), there are a number of technical (biology) related terms which are integrated into the lexicon of Oromo by the modern education. The loanwords under (117), passed a number of phonological adaptation during the integration process. Accordingly, there are also replacement, insertion and addition of vowels in the process of borrowing words. The impermissible initial consonant cluster is adapted to Oromo phonology through inserting the vowel *i* in between the clusters as in *pirootiinii* 'protein' and *iskaarvii* 'scurvy'. After changing the non-existing vowels, the English triphthong *æɪə* is replaced by *ayi* through inserting the glide *y* as in *ayireenii* 'iron'. Here, we can also note that the Oromo plural marker *-oota* replaced the English plural morpheme *-s* as seen in *orgaanizimoota* 'organisms' and *seeloota* 'cells'.

There are also other several scientific terms that are related to the field of chemistry which are loanwords of English. The following data are collected from grade seven chemistry textbooks to show the inclusion of English loanwords in Oromo.

(118) English	Oromo	adaptation (replacement/insertion/addition)
bromine /broʊmi:n/	→ <i>biroominii</i>	/oʊ/→/oo/, /ii/ addition
carbon /kɑ:rbən/	→ <i>kaarboonii</i>	/ɑ:/→/aa/, /ə/→/oo/, /ii/ addition
chlorine /klɔ:ri:n/	→ <i>kilooriinii</i>	/ɔ:/→/oo/, /ii/ addition /i/ insertion
cooper/kɑ:pər/	→ <i>koopparii</i>	/ɑ:/→/oo/, /ə/→/a/, /ii/addition
hydrogen/hɑɪdrədʒən/	→ <i>haayidiroojinii</i>	/aɪ/→/aayi/, /ə/→/oo/ə/→/i/, /ii/ addition
iodine /aɪdaən/	→ <i>ayoodinii</i>	/aɪ/→/ayoo/, /ə/→/i/, /ii/ addition
mercury /mɜ:kjəri/	→ <i>meerkurii</i>	/ɜ:/→/ee/, /jə/→/u/, /ii/ addition
oxide /ɑ:ksaɪd/	→ <i>oksaayidii</i>	/ɑ:/→/o/, /aɪ/→/aayi/, /ii/ addition
oxygen /ɑ:ksɪdʒən/	→ <i>oksijinii</i>	/ɑ:/→/o/, /ɪ/→/i/, /ə/→/i/ /ii/ addition
potassium /pətəsiəm/	→ <i>pootaasiyemii</i>	/ə/→/oo/, /æ/→/aa/, /iə/→/iye/, /ii/ addition
sodium /soʊdiəm/	→ <i>soodiyeemii</i>	/oʊ/→/oo/, /iə/→/ee/, /ii/ addition

As can be seen from the above data, several scientific terms that are related to the field of chemistry entered into Oromo with modifications. The loanwords under (118) are chemical elements related terms which are integrated into the lexicon of Oromo during textbook preparation. There are several vowel adaptation processes as it depicted in (118). One of such process is the replacement of the English diphthong *oo* into either /oo/ or /o/ as in *soodiyeemii* 'sodium' /soʊdiəm/ and *ovaarii* 'ovary' /oʊvəri/ as seen in (118) and (117) respectively. Another phonological adjustment seen in (118) is the insertion of *i* between the two consonants in order to avoid initial consonant cluster as in *biroominii* 'bromine' and *kilooriinii* 'chlorine'. This is due to the fact that Oromo disallows word initial consonant cluster. Oromo allows words to end with a consonant as in, for instance, *kun* 'this'; almost all the loanwords of source language (English) ending with consonant sounds, however, are adjusted via adding, most often the time, *ii* as in *soodiyeemii* 'sodium' /soʊdiəm/ and in most of the data we have seen so far. There are also cases in which the adjustment is made with the terminal vowel *a* as in, for instance, *giraama* 'gram' /græm/ and all loanwords under (116).

5.2.1.3.1.1.3 Arabic loanwords in the textbook corpus

Oromo is one of the languages of Ethiopia, which is found to enrich its lexicon through borrowing words from other exogenous languages such as Arabic. There are several loanwords from Arabic that are introduced into the vocabulary of Oromo. The data in (119) present some of Arabic loanwords which are collected from grade six, ten and eleven Oromo, grade seven and eight biology and civics, as well as from grade four environmental science textbooks.

(119) Oromo	gloss
<i>ajandaa</i>	'agenda'
<i>jabanaa</i>	'jug or kettle of coffee pot'
<i>kaartooniicha</i>	'the carton'
<i>kiraa</i>	'rent'
<i>maqasii</i>	'scissor'
<i>muuzii</i>	'banana'
<i>qalbi</i>	'grace'
<i>sa'aatii</i>	'clock'
<i>shukkaara</i>	'suger'
<i>suuqiiwwan</i>	'shops'
<i>zayitummaa</i>	'oilness'

Looking into the above loanwords that are incorporated into Oromo, even the native speakers of the language may wonder whether loanwords under (119) are really loanwords or the endogenous words of Oromo. But, after making deep investigation, any researcher or the native speakers of the language themselves can reach on to the fact that these words are Arabic loanwords. Here, it is some how difficult to clearly put how these loanwords entered into Oromo lexicon as there are at least two options for the loanwords to be introduced into the language. These loanwords might have entered into Oromo indirectly from Amharic as there was a trade with Arabs such as Yemens and Amharic speakers (Regourd & Nancy Um 2017:54). In addition, almost all of the above loanwords that are introduced into Oromo are also exist in Amharic too. On the other hand, these loanwords might have introduced into the language as Oromo speakers them selves have a direct contact (just like the existence of direct cultural contacts between

Arabs and Hadiyya people as stated in Samuel 2017:224) in the western part of the country. The people living in the eastern, southeastern and some south-western part of the country are Moslems and such loanwords might have entered into the language during religious expansion. Here, what can exactly be determined is that these Arabic loanwords are integrated into the lexicon of Oromo and they are included in the various textbooks of Oromo during textbooks preparation by textbook writers. These loanwords are now being taught to the students of Oromo as the lexicon of the language.

We can also note from the data in (119) that loanwords take the grammatical features of the recipient language (Oromo) and not the source language's grammatical features. For instance, the definite and plural marker which is attached at the end of loanwords as in *kaartooniicha* 'the carton' and *suuqiwwan* 'shops' respectively, clearly indicate that loanwords that are introduced into the language can only follow the grammatical features of Oromo. That is why the inflectional morphemes of the recipient language that we have seen above are attached at the end of the loanwords. Looking into the data in (119), one can also note that Arabic loanwords also seem to undergo a modification of derivational morphological structure to achieve harmony with morphological systems of the recipient language (Oromo) as in *zayitummaa* 'oiliness' in which *zayitii* is 'oil' and *-ummaa* '-ness' which is derivational morpheme.

5.2.1.3.1.4 Italian loanwords in textbook corpus

Oromo is also one of the Ethiopian languages that borrowed a number of words from other non-endogenous language such as Italian. The following data shows some of Italian loanwords which are incorporated into the lexicon of Oromo by the textbook writers. The data in (120) are taken from the textbook corpus (grade ten & eleven Oromo, grade seven biology as well as from grade one and four environmental science textbooks).

(120) Oromo	gloss
<i>gaazeexaa</i>	'newspaper'
<i>kaartaa</i>	'carta'
<i>kaartoonii</i>	'carton'
<i>paastaa</i>	'pasta'
<i>poostaa</i>	'posta'
<i>salaaxaa</i>	'salad'
<i>simintoo</i>	'cement'
<i>xaasaa</i>	'can'

Textbook writers also used Italian loanwords to cover the lexical gaps due to the non-existing concepts in Oromo. Italian is, therefore, one of the non-indigenous languages that become the source of loanwords for Oromo. Arguably, the Italian loanwords might have introduced into the lexicon of Oromo indirectly via Amharic as there was no direct contact between the Italian and Oromo speakers. Rather, there was language contact between Amharic and Italian as a result of occupation (Samuel 2017:193; Takkele 2000:91). There is, obviously, a phonological adjustment while loanwords enter into a language. There are terminal vowel lengths in all loanwords as seen in (120). Word medial length is also attested in all loanwords except the case of one loanword (i.e., *simintoo* 'cement').

5.2.1.3.1.1.5 French loanwords in the textbook corpus

Textbook writers also used French as a source of loanword in order to express previously non existing concepts in Oromo. The following loanwords of French are taken from grade one and four environmental science, grade two mathematics and grade eight Oromo textbooks.

(121) Oromo	gloss
<i>baabura</i>	'train'
<i>biiroo</i>	'bureau'
<i>foolisii/poolisii</i>	'police'
<i>kootii</i>	'coat'
<i>miiliyoona</i>	'million'

<i>ripaabilikii</i>	'republic'
<i>saantima</i>	'cent'
<i>shamiizii</i>	'shirt'
<i>sijaaraa</i>	'cigarette'

The above data shows the existence of French loanwords in textbooks. French loanwords are integrated into the teaching materials during the textbook preparation. The above loanwords of French seem to exist in both Amharic and Oromo. Here, just like other exogenous sources of borrowing, it could be assumed that the above loanwords of French were introduced into Oromo via Amharic as Amharic had got a higher status and was serving as a linguafranka, medium of instruction and administration and so on.

Generally, it has been noted that Oromo tried to express certain foreign concepts via searching from its own vocabulary before directly going into the borrowing option. The language has also enriched itself via borrowing from both endogenous and exogenous languages. Based on the principles of incorporating foreign concepts into a language, it is obvious that borrowing first from endogenous languages is preferable than directly taking loanwords from exogenous language. Contrary to the principles, of these two, exogenous languages were the major sources of enriching Oromo as compared to endogenous as seen in the above subsections.

From the Ethiopian language source of loanword point of view, Oromo preferred Amharic to incorporate loanwords more than any other Ethiopian languages. Amharic, a language which has been serving as a linguafranka in Ethiopia for several years, become the major endogenous language for the enrichment of the lexicon of Oromo as there are several Amharic loanwords than any other indigenous languages in the country. This fact can be attested via looking into the number of loanwords collected from the textbooks. The direct socio-cultural contact is also one of the possible reasons for loanwords to exist in the lexicon of Oromo. It is also possible to perceive that the bilingual Oromo textbook writers have played a role in integrating Amharic loanwords into Oromo. French and Italy loanwords seem to be introduced into the Oromo lexicon via an intermediary language (Amharic). There are only few loanwords of Greek (such as *maaraatoonii* 'marathon' and *Itoophiyaa* 'Ethiopia' which are introduced into Oromo via

Amharic. Although such loanwords have been included in the textbooks during textbook preparation, it does not mean that they entered into the language only during such event. This few Greek loanwords existed in the Amharic even before the expansion of regular education.

As far as borrowing in Oromo is concerned, English is attested to be the most preferable exogenous source language to enrich Oromo particularly for the designation of scientific and technological concepts. Several English loanwords are integrated into the primary and secondary school textbooks by the textbook writers. This fact can be proven via looking at the data collected from the current working textbooks and the textbooks at large (see Appendix E). The English loanwords in Oromo belong to various disciplines such as mathematics, chemistry, biology etc. Just like the case of loanwords of Amharic, there are also two perceptions as how English loanwords introduced into Oromo at large and in the textbooks in particular. The first perception, at least from the present researcher's point of view, is that an intermediary language (Amharic) could have played its role in transferring some of the English loanwords into Oromo via due to the fact that Amharic has been language of media, administration and medium of instruction perhaps before any of Ethiopian languages. This means English loanwords first might have entered into the lexicon of Amharic and then transferred into Oromo through socio-cultural contact. The second perception is that educated Oromo bilingual might have played their own role in integration processes of some of loanwords of English as the textbook developers are expected to be educated.

Several phonological adaptations took place during the inclusion of loanwords into the lexicon of Oromo. Here, what is more interesting is find out whether there are phonological adaptation during loanword integration from different sources into the lexicon of Oromo at the time textbook preparation. According, as per all the data collected show, there is no unmodified borrowing; all loanwords undergone modification. The non-existing vowels of the source languages are replaced with the nearest vowels of the recipient language (Oromo). On the other hand, there is no diphthong and triphthong in Oromo. Hence, the simple vowels (short or long vowels) or glide insertion will replace the diphthongs and triphthongs of English.

Apart from the vowel adaptation of loanwords, there is also some consonant adaptation. The non-existing English consonant /ŋ/ is seen to be replaced by the nearest consonant /n/ as in, for instance *reeктаangilii* 'rectangle' /rektæŋgl/ during the integration process. Furthermore, some previously non-native consonants like /ʒ/, /p/, /v/ are introduced into Oromo during textbook preparations by writers as a result of the integration of enormous scientific and technological terms in the textbooks. Vowel insertion is also one of the adaptation processes that took place during the loanword integration process. Accordingly the impermissible consonant cluster (two consonant cluster at the initial and three consonant clusters anywhere) is adjusted to the Oromo phonology via inserting either *i* and in very rare cases *o* as in, for instance, *pirootiinii* 'protein' /prooti:n/ and *rohombosii* 'rhombus' /ra:mbəs/ respectively. In some of the cases of the insertion of *o* to conform to Oromo phonology could be avoided by deleting *h* as in *roombosii*. Though Oromo phonology allows words to end with a consonant as in *afaan* 'mouth/ language', almost all loanwords in Oromo receive the terminal vowel /ii/ and /a/ as in, for example, *ayoodinii* 'iodine' /aidaən/ and *tiraktera* 'tractor' /træktər/ respectively.

5.2.1.3.1.2 Superfluous borrowing and parallel use of terms

Textbook writers have made unnecessary borrowing during textbook preparation processes despite the existence of genuine words in Oromo. Unnecessary loanwords are taken from endogenous and exogenous languages. The data in (122) are collected from grade one environmental science, grade eight mathematics, grade four environmental science textbooks.

(122) Loanword	source language	genuine word	gloss
<i>dinnicha</i>	Amharic	<i>moose</i>	'potato'
<i>gindii</i>	Amharic	<i>nooyyee</i>	'block'
<i>kooffiyyaa</i>	Amharic	<i>ginbaabee</i>	'cape'
<i>maqasii</i>	Arabic	<i>qaxxaamuraa</i>	'scissor'
<i>niyuutireentoota</i>	English	<i>dhangaalee</i>	'nutrient'
<i>parsantii (%)</i>	English	<i>dhibbantaa</i>	'percent'
<i>piroobleemii</i>	English	<i>gaafura</i>	'problem'
<i>qabattoo</i>	Amharic	<i>saqqii</i>	'belt'

<i>reeshoo</i>	English	<i>gahee</i>	'ratio'
<i>suqii</i>	Arabic	<i>dukkaana</i>	'shop'
<i>verteksii</i>	English	<i>fiixee</i>	'vertex'
<i>zeeroo</i>	English	<i>duwwaa</i>	'zero'

As can be seen from the data in (122), there are several unnecessarily borrowed words in Oromo even though genuine words of the language exist. Here, it can be noted that there are more English loanwords which are superfluous than any other languages as can be seen in (122).

There are several native terms of Oromo which are alternatively used in the textbooks as it is depicted in (123). The data are collected from grade three, eleven, twelve Oromo, grade eight biology and civics, grade seven biology and chemistry, grade three and six mathematics textbooks showing the use of alternative native terms of the Oromo in textbooks.

(123) Oromo word (main)	Oromo word (alternative)	gloss
<i>afoola</i>	<i>ogafaan</i>	'oral literature'
<i>balaa</i>	<i>miidhaa</i>	'accident'
<i>boohaarsaa</i>	<i>bashannanaa</i>	'impressionist'
<i>bu'aawwan</i>	<i>firiwwan</i>	'outcomes'
<i>buqqee</i>	<i>wullee</i>	'calabash'
<i>cimdii</i>	<i>xundaa</i>	'couple'
<i>faallaa</i>	<i>masanuu</i>	'opposite'
<i>gataa</i>	<i>kosii</i>	'dung'
<i>guugni</i>	<i>ogeessi</i>	'expert'
<i>hanqina</i>	<i>mudoo</i>	'gap'
<i>hurrii</i>	<i>duumessa</i>	'cloud'
<i>makaa</i>	<i>bulbula</i>	'solution'
<i>maxinoo</i>	<i>qixxaa</i>	'temper'
<i>qalaasa</i>	<i>xaasaa</i>	'can'
<i>qarabaa</i>	<i>haadduu</i>	'knife'
<i>qoosaa</i>	<i>baacoo</i>	'comedy'

<i>quxisuu</i>	<i>eegoo</i>	'junior'
<i>siiqfamu</i>	<i>butamu</i>	'pull'
<i>wadaroo</i>	<i>funyoo</i>	'rope'

As can be seen from (123), there are several instances of the use of alternative native Oromo terms in the textbooks that refer to the same concept. The practice of alternative use of native terms of a language has something to do with language standardization as the use of the standard language is believed to be strictly obeyed in the learning materials. The use of native terms to designate concepts is preferable before directly going to borrowing of words. The use of native words alternatively in the textbooks is arguable. The inclusion of alternative native terms (varieties) of a language in textbooks could be considered as ways of accommodating varieties of a language and their respective speakers. For instance, Wondimu (2015:365) noted that the varieties of Oromo should be represented in the textbooks via putting optional words in bracket. Wondimu (2015:66) further argues that the inclusion of varieties of a language should be encouraged via putting parallel words in parenthesis in school textbooks to show the richness of the language. On the other hand, Almaz (2018:124) argues that putting a word of a variety of a language in parenthesis next to another word could add difficulty instead of clarifying the concept to the learner. The author further argues that the existence of several parallel forms or words in the school textbooks can create a less stable norm and negatively affect the standardization process of a language.

Here, one can note that Wondimu (2015) seems to give less attention for the standardization of Oromo as compared to the inclusion of varieties of the language in textbook which have a fundamental effect to the development of the language. The present research does not pretend to argue against Wondimu's argument of inclusion of varieties of Oromo in the textbooks through putting alternative words of varieties of language in parenthesis. Further, this research did not argue that the use of native alternative words should be completely avoided. It is obvious that in some of the cases words could be used interchangeably as a means of synonymous expressions in textbooks. Rather, the over appearance of several parallel native words (as it has been attested in the textbooks of Oromo) which are taken from various varieties of Oromo definitely pull the standardization process back. This may also be a burden to the learners. The inclusion of

varieties of the language in the textbooks could be attained via consistent application of the codified words for each concept. Otherwise, the existence of several parallel native words written in parenthesis page after page may, even destandardize the language. To avoid such lexical variation; the standardization committee of Oromo is normalizing the language. But, as it has been stated in the focus group discussion session, the collaboration of the committee and the textbook writers are not to the expected level. This could be one of the reasons for the existence of several standardization issues in textbooks.

Apart from the use of parallel words in parenthesis, textbook writers have also opted to use one form (concept) in some instances and other forms in other instances for the same concept. As a result, two and more than two forms (words) are used for the same concept in several parts of the textbooks. The following data are collected from grade five social science, grade seven & eight and from grade eleven Oromo textbooks.

(124) Oromo word (main)	Oromo word (main)	gloss
<i>aangoo</i>	<i>baallii</i>	'authority'
<i>baallii</i>	<i>aangoo</i>	'authority'
<i>bokkaa</i>	<i>roobaa (coba)</i>	'rain'
<i>bushaa</i>	<i>xiqqaa</i>	'small'
<i>gabatee</i>	<i>naanniga</i>	'trough'
<i>goolaba</i>	<i>xumura</i>	'end'
<i>maasii</i>	<i>oyiruu</i>	'farm'
<i>naanniga</i>	<i>gabatee</i>	'trough'
<i>oyiruu</i>	<i>maasii</i>	'farm'
<i>roobaa</i>	<i>bokkaa (coba)</i>	'rain'
<i>xiqqaa</i>	<i>bushaa</i>	'small'
<i>xumura</i>	<i>goolaba</i>	'end'

As can be seen from (124), two or more words are used as a main word; but not as alternative words which are expressed in parenthesis as we have seen in (123). The expression of a concept in one form from one variety at one time and using another form at another time in a language

being standardized will make the process stagnant. As can be seen from (115), there are several instances of writing the same concept in textbooks with various forms from varieties of Oromo. It is obvious that when a language is at its early stage of standardization, representing the same concept with more than two words may be acceptable. But, standardization of Oromo may not be leveled at the beginning of its process as the standardization committee of Oromo has been working for at least more than twenty five years; even though what has been normalized is so little as compared to what is left. Hence, observing such kind of variation in a site where higher regularization of language is practiced is totally unacceptable. This could be due to the result of loose of coordination between the standardization committee of Oromo and education authority in general and textbook writers in particular as can be seen from the discussion during the FGD3 with the committee (see Appendix G, FGD 3, P3, 20). All in all, the use of words inconsistently in the textbooks has a potential to create irregular forms and pull back the standardization process.

Apart from the use of Oromo parallel terms within itself, textbook writers also used English loanwords in parallel with Oromo words. The following data are taken from grade eight biology, grade seven and eight chemistry, grade eight mathematics as well as from grade three, eleven and twelve Oromo textbooks for an illustration of the use of parallel terms in the textbooks.

(125) Parallel terms (Oromo and English loanwords)	gloss
a) <i>gabatee / chaartii</i>	'chart'
<i>hammamii / daameenshinii</i>	'dimension'
<i>qarxiwwan / kuwaadraantota</i>	'quadrants'
<i>saphuuphuu / algeen</i>	'algae'
b) <i>eerzii / dachee</i>	'earth'
<i>ko'oordineetii / cimtii tartii</i>	'coordinate'
c) <i>antima / objaktii</i>	'object'
d) <i>addunyummaa / gilobaalizeshinii</i>	'globalization'
<i>akkayyaawwa / modeelota</i>	'models'
<i>dacha'oo / sameetrikii</i>	'symmetric'
<i>danguleessa / piinkii</i>	'pink'

<i>ho'inaa/tempireecheerii</i>	'temperature'
e) <i>afaala daree/filoorii</i>	'floor'
<i>dhuubbaa zayitaa/ peetirooliiyemii</i>	'petroleum'
<i>gabatee gurraacha/bilaak boordi</i>	'black board'
<i>kal-heddee /maltidaamenshiinalii</i>	'multidimensional'
<i>leensiilamee /baayinookulaara</i>	'binocular'
<i>ujummoon sanyii dhiiraa /vaasdiffaransii</i>	'vas deferens'
<i>wantoota jeneetikii /nuukilooyidii</i>	'nucleoid'

From the data presented in (125), we can note that most of the first terms used as parallel terms are Oromo native terms. The native terms used in parallel with loanwords are derived as in (125d) (*addunyummaa* 'globalization' (*addunya* 'world'+ *-umma*'), blends as in (125c) *antima* 'object', simplex words as in (125a) *gabatee* 'chart' and compounds as in (125e) *gabatee gurraacha* 'black board'. There are more interesting issues in (125) regarding the parallel use of terms. Textbook writers also used some initial loanwords in parallel with simplex and compound words as in (125b) to express foreign concepts. In some case textbook writers also used words and compounds which are formed via loan translation method to use them parallel with loanwords as in, for instance, *kal-heddee /maltidaamenshiinalii* 'multidimensional'. Compounds that are translated using the techniques loan translation + loanword and native term + loanwords are also seen to be written parallel with loanwords as in, for example, *leensii lamee /baayinookulaara* 'binocular' and *wantoota jeneetikii /nuukilooyidii* 'nucleoid'

Looking into the parallel use of terms in (125), one can note that loanwords are being used in the textbooks as a main word in the existence of Oromo native equivalent words as in (125b), for instance *eerzii/ dachee* 'earth'. This will certainly affect the standardization of the language as the first option is searching a native word to express a concept before looking for loanwords. The use of native could have been more appropriate for the transparency of the meaning to the learner. To the contrary, the use of short loanwords could be more important than using very long compounds and phrases, particularly when one wants to achieve the principle of word economy as in *ujummoon sanyii dhiiraa /vaasdiffaransii* 'vas deferens' (see also Samuel 2017:190). Such situations of use of native terms of Oromo parallel with loanwords may delay the progress of

standardization process of Oromo as it pulls back the use of technological and terms that are related to a particular field.

5.2.1.3.1.3 Summary of borrowing in Oromo

Borrowing is one of the lexical elaboration techniques which used extensively to enrich Oromo. There are several scientific and technical terms which are borrowed and included in the textbooks and in other texts of Oromo. The general pattern of the source languages which are used to enrich the lexicon of Oromo is summarized based on the data collected from textbook corpus and from the materials published by the committee (see also appendix E for lists of source languages and loanwords). The summary of the source languages is presented in table (20) as follows.

Table 20: Summary of loanwords in Oromo in general (see appendix E)

Language	Frequency of loanwords	Percent
English	101	52.4%
Amharic	33	17.1%
Arabic	32	16.6%
Italy	11	5.7%
Swahili	7	3.6%
French	7	3.6%
Greek	2	1%
Total	193	100%

As can be seen from the summary of the loanwords presented in the above table, seven languages become the source languages in the process of elaborating Oromo. Of these, to the best of access of data sources used, there is only one Ethiopian language (i.e. Amharic) that becomes the sources for enriching the lexical of Oromo. This language is also an intermediary language to transfer loanwords from other languages. We can also note from the summary of loanword that, except Amharic, based on the sources of data for the present study, other Ethiopian languages are not used as source languages. This is, however, does not mean that there

are no other Ethiopian languages which became the sources of lexical elaboration in various texts beyond the scope of this study. As per the summarized data, most of the source languages which are used to elaborate Oromo are foreign languages. Accordingly, English takes the lion shares of the foreign source of languages followed by Arabic. English plays paramount roles in becoming the exogenous source language, particularly, for the inclusion of scientific and technical terms of foreign concepts. As can be clearly noted from the summary of source languages stated in table 20 above, Greek is the least preferred source language to expand the lexicon of Oromo.

Looking into the whole pattern of lexical elaboration of Oromo, we can note that the language used its internal means extensively to express foreign concepts before directly going into the act of borrowing from other languages. The indigenous language (Amharic) is found to be the second (next to the internal means) option of borrowing in the course of the enrichment of Oromo. As far as borrowing is concerned, all borrowed words undergo modifications to fit to the phonology of the language. Hence, the present study did not find unmodified borrowing as in the case of Amharic borrowings as stated in Samuel (2017:183). We can also note from the summary in table (20) and appendix E that the exogenous languages are the third option to become the source for elaborating the lexicon of Oromo.

5.2.1.3.2 Loan translation in Oromo textbooks

Apart from the direct borrowing that we have seen in the previous sections, textbook writers elaborated the lexicon of the language through indirect borrowing. One of such indirect borrowing is loan translation. In loan translation is the borrowing of meanings from a donor language and then incorporating in the lexicon of the recipient language (Wach 2013:161). In such technique of lexical enrichment, the word forms of the donor language will not be borrowed. On the other hand, a loanword could be characterized as a word which is borrowed from a donor language and integrated into a recipient language with the word form. Word of the donor language will not be translated into the recipient language in this case.

Loan translation can be classified into literal loan translation and conceptual loan translation. Literal translation is regarded as the closest possible rendering of a foreign language text into another language and is a translation in which “the source language grammatical constructions

are converted to their nearest translation language equivalents, but the lexical words are again translated singly, out of context” (Lomaka 2017:235). The author further states that it is understood as a kind of translation that characterizes the way of rendering the source text and the approach to achieving equivalence. This suggests that all elements of the source text are reproduced one by one without regard to their position and possible changes of their meaning in the context, to the detriment of their communicative significance, which often leads to the distortion of all levels of the text (Lomaka 2017:238).

Non-literal meaning is the sharply distinguished complement of literal meaning. It is considered pragmatic (extra linguistic), it is associated with the utterance and the speaker (rather than the sentence), and it is non-conventional and non-compositional (Ariel 2002: 361). Non-literal meaning is indirect. It is derived by combining literal meaning and general cognitive inferential processes (Ariel 2002:361). The literal loan and non-literal (conceptual) loan translation in Oromo textbooks are presented in in following section.

5.2.1.3.2.1 Literal loan translation in Oromo textbooks

There are several instances of literal loan translation in the textbooks that are used by the textbook writers to cover the existing lexical gap. The following data are collected from grade twelve Oromo, grade eight biology and mathematics as well as from grade seven and eight social study textbooks.

(126) English term	literal loan translation	gloss
altitude	<i>olka'iinsa lafaa</i> up altitude earth of	'up altitude of'
cold war	<i>waraana ololaa</i> war speech of	'war of cold'
commutative property	<i>amala jijjirraa</i> property distribution of	'property of commutative'
distributive property	<i>amalaraabsamaa</i> property distributive of	'property of distribution'
imperfective aspect	<i>kanhin-raaww-at-am-n-e</i> GN NEG-complete-MD-PS-NEG.IMP	'imperfective aspect'

mouth of river	<i>galalagaa</i> enterance of river	'of river mouth'
perfective	<i>raawwatame</i> done	'perfective'
salty water	<i>bishaan soogiddaa'oo</i> water of salty	'water of salty'
subtropics	<i>hoo'a walakkeessoo</i> temperature of half	'subtropics'
symmetric	<i>dacha'oo</i> symmetric	'symmetric'

As can be seen in (126), textbook writers used several literal loan translations that have the structure of words and compounds so as to express the newly entering concepts in Oromo from various spheres (political, economic and social). Looking into the data in (126), one can note that there is structural overlap of compounds (an internal means of lexical elaboration) and literal loan translation (external means of lexical elaboration i.e. meaning borrowing). For instance, *bishaan soogiddaa'oo* 'salty water' fulfils both criteria of literal loan translation (as it is a direct translation of both the form and intended meaning of English words into Oromo) and compound (as the structure much with structure of compounds). The English words such as 'constant' are also seen to be used in the textbooks via translated in its Oromo equivalent word called *dhaabataa*. Apart from the overlaps of compounds and literal loan translation, there exist also overlap of derived words and literally translated words as in, for instance, *dacha'oo* 'symmetric' since *dacha'oo* is derived from the original word of Oromo called *dacha* 'double'.

5.2.1.3.2.2. Non literal loan translation in Oromo textbooks

Apart from the literal loan translation, English terms are also seen to be translated non-literally (conceptually) by textbook writers in order to cover the existing lexical gap while expressing new concepts entering the language. The following data collected from grade eight mathematics and biology textbook.

(127) Term	non-literal loan translation	gloss
coefficient	<i>maxxantuu</i>	'sticky'
column	<i>irraan gadee</i> top to down	'top to down'
equation	<i>hima wal-qixaa</i> statement equal of	'statement of equal'
equivalent equation	<i>himoota wal-caalmaa</i> statement not equal of	'statement of equivalent'
expanded	<i>diddiriirsuu</i> 'to stretch'	'to stretch'
higher common multiple	<i>hirmaata walii guddicha</i> participant common of higher	'higher common multiple'
igneous rock	<i>dhagaa baq-qorrii</i> stone melted cold of	'cold stone melted of'
metamorphic rock	<i>dhagaa jijjiiramaa</i> stone changed of	'changed stone'
right angle triangle	<i>rog -sadee kofa sirrii</i> side three of angle right	'right angle of three'
row	<i>toora dalgaa</i> to horizontal	'horizontal'
<i>seerviiksii</i>	<i>afaan gadameessaa</i> mouse uterus of	'mouse of uterus'
simplify	<i>salphisi</i> simplify	'simplify'
temperature	<i>haala hoo'aa-qabbanaa</i> situation hot and cold of	'situation of hot and cold'

As seen in (127), there are a number of words, compounds and phrases that are translated non-literally by textbook writers so as to express foreign concepts. Hence, looking into the textbooks in general and into (127) in particular, one can conclude that non-literal loan translation is extensively used in the education sector to develop the lexicon of the language. Eventhough non-

literally translated words, compounds and phrases are better than direct borrowing both forms and meanings, translating foreign concepts with by far very long words or compounds or phrases as compared to source language's words, compounds or phrases which deny the principle of economy. For instance, a very long expression *haala hoo'aa-qabbana* is translated non-literal to represent a very short English term 'temperature'. This is totally unnecessary long and not economical expression which becomes a burden for the learners.

On the other hand, the source language for all of the non-literal translation in (127) is English. Technical terms are non-literally translated to cover the lexical gap via native words or compounds or phrases. Most of the technical terms which are translated non-literally in (127) belong to mathematics field of study. But, this does not mean that there are no other non-literally translated words, compounds and phrases from other else field of studies.

Oromo textbook writers have extensively used a mixed approach (literal and non-literal loan translation with loanwords) to designate foreign concepts in Oromo textbooks. The following data collected from grade eight, biology, chemistry, mathematics, civics and ethical education, grade four environmental science, grade seven social study textbooks clearly show the fact mentioned earlier.

(128) terms	integrated words (compounds)	method integration
anti-corruption commission	<i>komishinii farra malammaltummaa</i>	loan + loan translation
electric power	<i>anniisaa elektirikii</i>	loan translation + loan
electric translation	<i>dabarsoo elektirikii</i>	loan translation + loan
human right commission	<i>komshiinii mirga namumma</i>	loan + loan translation
perfect square	<i>iskuweeri isirrii</i>	loan + loan translation
physical change	<i>jijjiirama fiizikaalaa</i>	loan translation + loan
red blood cells	<i>seeloota dhiigaa diimaa</i>	loan + loan translation
two type atom compound	<i>kompaawundii atoom-gos-lamee</i>	loan + loan translation
unicellular organism	<i>orgaanizimoota seel-qaenxee</i>	loan + loan translation
universal scale	<i>iskeelii maralaalu</i>	loan + loan translation
world bank	<i>baankii addunyaa</i>	loan + loan translation

As can be noted from (128), textbook writers have used the combined approach to designate scientific and technological terms in addition to the independent use of the direct borrowing and loan translation. For instance, the foreign expression 'anti-corruption commission' *komishinii farra malammaltummaa* is integrated in Oromo via combining both direct borrowing of the word *komishinii* 'commission' and through loan translation of the foreign compound 'anti-corruption' *farra malaammaltummaa*. Such way of integrating new concepts into Oromo has been attested in the textbooks at several pages. Hence, be it in terms of direct borrowing or indirect borrowing (i.e. borrowing of meaning; but not the form as in the case of loan translation), an exogenous language (English) played a paramount role than any other source languages in the course of the elaboration (particularly in the lexical enrichment) of the language in particular and in the development of the language in general.

5.3. Comparison between lexical elaboration by the committee and textbook writers

Both the standardization committee of Oromo and textbook writers elaborated the lexicon of Oromo via both internal and external means. Meaning extension, redefinition of meanings, derivation, abbreviation, blending and compounding are among the internal means which both the standardization committee of Oromo and textbook writers used to enrich the Oromo. Of these internal means of lexical elaboration, meaning extension, derivation, compounding, blending and abbreviation are extensively applied by textbook writers than the standardization committee of Oromo. Derivation is also extensively used in the textbook and in the materials compiled by the standardization committee of Oromo to elaborate Oromo. Regarding external means of lexical elaboration, borrowing and loan translation are used to elaborate the lexicon of the language in both arenas. Several scientific and technical terms are included in the lexicon of Oromo via extensively using both borrowing and loan translation in the textbooks by the textbook writers more than by the standardization committee of Oromo. All loan words in Oromo undergone phonological adaptation process and unmodified borrowing is not attested in all of the documents (in the publications of both the committee and textbooks).

There are orthographic inconsistencies in publications of both the standardization committee of Oromo and educational materials (primary and secondary Oromo). There are no rules

(regulations) in any document of the standardization committee of Oromo as to how compounds should be spelled. Textbook writers and even the standardization committee of Oromo itself spell compounds in any way they want. In some of the cases compounds are written with hyphens and in another case with white space between the constituents of the compounds. Still in some other cases the same compound is spelled as a single word. Unlike Oromo in which there are no officially stated rule to write compounds, orthographic rules are clearly stated to write compounds in Hadiyya and hence, the use of a hyphen is not suggested and there are only two options to write compounds; as a single word and with space in between the words (Samuel 2017:171). Both the standardization committee of Oromo and textbook writers used abbreviation in their courses of elaborating the language independently (on their own ways). The method, however, is extensively used during textbook preparation by the textbook writers than by the committee. There are also variations in writing abbreviations in both documents as there are no codified rules in order to consistently abbreviate words.

Complex compounding strategy is also applied to elaborate the lexicon of Oromo in documents compiled by the standardization of Oromo and textbook developers. Complex compounding strategy having up to six constituents is extensively applied in the textbook corpus in order to express scientific and technical terms. But, the complex compounding method applied by the standardization committee of Oromo is less complex. Compounds that have no more than three constituents are applied by the committee and most of the compounds that exist in the publications of the standardization committee of Oromo are of simple compound having only two constituents.

Generally, although the standardization committee of Oromo and education sector has expected to do their own task, i.e., standardization (normalization) and implementation of the codified norms respectively, both of them are engaged in the lexical elaboration of the language. The research participants state during the focus group discussion that textbook writers are performing the task of the standardization committee of Oromo since scientific, technical and education related terminologies are not still codified by the committee (see appendix G, FGD3, 28). Writers are trying to cover the lexical gaps while preparing textbooks. The participant further explains that textbook preparation should involve standardization personnel, linguist, educational

expert and lexicographers. The language to be used to prepare school textbook is expected to be standard as students are expected to learn standard language. As opposed to this, though textbook compilers played their own role in elaborating the language during textbook preparation, there are still inconsistencies and variations in the textbooks. In addition, the education bureau of Oromia and the standardization committee of Oromo are not working hand in hand to the extent needed (Appendix G, FGD 3, 28).

5.4. Summary

In this chapter, I have described the role the education domain play in the process of (de)standardization of Oromo. Accordingly, though the task of education sector is to implement what have been codified (normalized) by an authoritative body (a language standardization committee in the case of Africa and language academy in European countries), this study identified that textbook writers are actual doing what the standardization committee of Oromo could have done. A number of lexical elaboration methods are employed in the textbooks to cover the existing lexical gaps. Textbook writers used both internal and external means to elaborate the language. Meaning extension, redefinition of words, derivation, compounding, blending and abbreviation were among the internal means which are employed to designate concepts in the education domain. Of the internal means of lexical elaboration, compounding, derivation, extension of meaning, blending and abbreviation are among the productively employed in the process of the lexical elaboration in the textbook corpus this study consulted.

Moreover, borrowing and loan translations are employed by the textbook writers to enrich the lexicon of Oromo. Both of them are extensively employed by textbook writers to cover concepts related to scientific and technical terms. Unnecessary (superfluous) borrowing is also seen in the textbooks while there are genuine words in the language. Spelling inconsistencies of compounds are also one of the challenges that may negatively affect the standardization process of the language. There are also lexical variations in the textbooks that need codifications. On the other hand, there are also challenges, from textbook writers' side, of putting the normalized forms into practice. This shows that education sector and the authoritative body (the standardization committee of Oromo) is not working hand in hand. One of the participants of the research during the focus group discussion stated that what is codified is not put into practice due to lack of

distribution of the codified items and lack of coordination among the standardization committee of Oromo and other stakeholders in general (see Appendix G, FGD 3, 28).

Generally, although there are problems in consistent application of forms in the textbooks, the education sector (the primary and secondary school textbook writers) is positively contributing to the standardization process of the language. The textbook writers played a paramount role, particularly in the lexical elaboration process of the language which the standardization committee of Oromo should have done. Textbook writers coined scientific and technical terms to designate new concepts.

CHAPTER SIX

SUMMARY, CONCLUSIONS AND RECOMMENDATION

6.0 Introduction

The last chapter of this dissertation summarizes the main issues discussed in the present study in general. This chapter of the dissertation consists three parts. Important activities included in the research are revised and presented in the summary section of this chapter. The result of the analysis of all data needs a conclusion and hence, is presented in the conclusion section of the same chapter. Finally, based on the findings of the study, recommendations about the standardization processes are forwarded under the recommendation section.

6.1 Summary

The present research was conducted with the aim of describing the orthographic and lexical standardization of Oromo. The research focused on orthographic and lexical perspective of standardization of the language. This is due to the fact that the standardization of orthography and lexicon of a language play a paramount role to the standardization of a language in general. As a result, the present study is envisioned to investigate the core issues of standardization of Oromo.

It is obvious that any language has to pass at least some sort of standardization before being introduced in both as a medium of instruction and as an independent subject at any level of education. In this regard, it is possible to regard Oromo as a language that is being standardized as it is being codified by the standardization committee of Oromo, used as a medium of instruction in the primary schools, being taught as an independent subject at all levels of education. It is also declared to be the official working language of the federal government of the country in this year (MoCT 2020:14). Hence, there is no argument that such placement of Oromo clearly shows that there is an urgent need for further tasks of standardization of the language.

The present study is aimed at answering four research questions: How can the orthography of Oromo be evaluated against the principles of orthography standardization? How is Oromo being codified by the standardization committee of the language? What are the linguistic techniques applied to elaborate the language? What are the challenges of standardization of Oromo? In

order to get possible answers to these research questions, data were collected via various qualitative data gathering techniques such as focus group discussions, document analysis and interview.

The study first dealt with the evaluation of orthography development and its standardization. The section starts with the description of the codification process and existing debate regarding the choice of script for writing Oromo. The section also analyzed the orthography of the language against principle of orthography. The main data used in the present study came from the materials compiled by the standardization committee of Oromo, from the current working school textbooks written in Oromo and from an interview and focus group discussions with members of the committee. The other main concerns of the present study were the identification of lexical elaboration techniques and the challenges of standardization of the language. The lexical elaboration of the language by the authoritative body (the standardization committee of Oromo) and the role the education sector play in the (de)standardization processes of the language is discussed in separate sections and are compared. Finally, based on the data collected data for the above thematic area, conclusions are drawn and possible recommendations are forwarded in the following subsections.

6.2 Conclusion

The following conclusions are drawn from the critical analysis of the data collected through the application of qualitative data collection techniques such as document analysis, focus group discussions and interview. The result of the analysis of data showed that Oromo was written using various scripts until the end of 1990. The study finds out that there are still arguments regarding the choice of scripts to write Oromo. The arguments are both for and against the use of Ge'ez and Raman based Oromo Script. The present research, however, argues for the use of Roman based script (*Qubee*) to write Oromo based on the linguistic, pedagogical, practical, acceptance and from the country's language policy point of views (see 3.3.2.1). Oromo was written using three scripts (Ge'ez, Arabic and Roman based scripts by various individual writers as there was no legally approved script to write Oromo. The Roman script (which is evolved from the Latin script) was officially accepted in 1991 to write the language. Developing orthography is a vital aspect of language standardization. In the same way, revisiting or

reconsidering and standardizing the developed orthography of a language is almost equally essential. Some of such issues are described in the previous chapters. In this regard, despite there are issues that require the need for revisiting the Roman based script of Oromo (*Qubee*), except the attempt made by the education bureau of Oromia, which was immediately reversed due to politics and acceptance issues, it is not yet officially reformed.

As it has been discussed in chapter three, <ts> /s'/ unnecessarily introduced into the phonemic inventory of the language (as per the decision of the standardization committee of Oromo) since it occurs in a few borrowed words like <tsabala> 'holly water' (see section 3.7.2). But this phoneme could have been written using the nearest grapheme <x> /t'/ as in <xabala>. The phoneme /z/ which is introduced into the language; but not included into both the phonemic inventory and the alphabet order of the language is being written using various graphemes like <zh>, <zy>, <zn> and <j>.

On the other hand, the unsettled issue about the graphematic representation of glottal stop or sound represented by /ʔ/ in the IPA is one of orthographic standardization problems discussed in the present study. It needs to be represented by the appropriate grapheme and normalization. Despite of its occurrence in the language and its inclusion as an independent phoneme in the phonemic inventory of the language, the grapheme for this phoneme is not still devised and included in orthography of the language. It is represented with a diacritic marker ('). This study recommends the glottal stop better be reconsidered by the standardization of Oromo for at least three reasons. First, it is a phoneme which is included in the phonemic inventory of the language by the standardization committee of Oromo and, hence, better to be represented with an appropriate grapheme rather than with punctuation marker ('). Second, it is not yet included in the alphabetic order of the language. Third, representing a phone with a diacritic is not recommended in the existence of graphemes as diacritic marks are extra marks used together with phonetic symbols in order to represent the actual pronunciation of phonemes (Stark 2010:91).

Hence, this study argues for the graphematic representation of the glottal stop with IPA symbol /ʔ/ and its usage in word internally only as all most all words begin with a consonant. All those

words that begin with a vowel in orthography of the language have a glottal stop at the beginning as in, for instance, *arraba* [ʔarraba] 'tongue'. Therefore, we can omit it uniformly as it is known to exist. Its place in the alphabetic order seems at the beginning. But, to avoid its confusion with the first alphabet <a>, it should be placed at the end of the alphabet order of the language.

There are also unsettled issues of underrepresentation of digraphs in their geminated form of realization. To this end, all phonemes represented by digraphs are being written using the non-geminated form in the current orthography of Oromo whether they are realized as geminated or as non-geminated. Since the issue of the representation of gemination has been one of the factors for the decision to adopt the Roman based script (*Qubee*) to write Oromo, it is vital that careful attention has to be given to it with a view to making it transparent, simple and more economical to use. Hence, there is a transparency problem regarding the marking of the geminated form of digraphs. This study, therefore, recommends that the geminated realization of these sounds represented by digraphs has to be marked. But this study argues that doubling the entire digraph (just like the case of Hadiyya) as in, for instance, <hodhdhu> 'to sew' can even worsen the previous problems of (unmarkedness) as doubling the entire digraph is not economical. As a result, this study recommends the doubling of only the first constituent of the digraphs as in <hoddhu> 'to sew' (see section 3.7.2). Concerning maximum transfer of skills, representation of the ejectives <x> [tʰ], <q> [kʰ], <c> [tʃʰ] and <ph> [pʰ] may have a negative impact on the transfer of reading skills in English and Oromo; particularly at the early stage of learning (see section 3.7.2). Of course, an experimental study has to be done to strengthen the issue.

Another important point that needs revisiting is the alphabetical order of the language. The alphabetical arrangement of the graphemes of Oromo is also one of orthographic standardization issues that need further consideration. The current alphabetic alignment of characters of the language is not systematic, particularly with regard to the order of digraphs in the alphabet. The digraphs are ordered separately from the monographs. This may negatively affect learning. In addition, the frequently occurring glottal stop has not been represented with grapheme and included in the alphabetic order of the language. The present study suggested a better arrangement of alphabetic order under the third chapter of this study to be reconsidered by the concerned authority (see section 3.8 and Appendix C).

There exists also standardization gap regarding the normalization of Oromo ordinal numbers as there are various options of representing it. Accordingly, *-eeysaal/-eeysool/-eessaa /-eessool/-eechaa/-eechoo/ -affaa/-affoo/* are the possible representation of ordinal numerals in Oromo. One of these ordinals has to be codified and normalized as there are eight ways of representing Oromo ordinals. The present study argues for *-affaa* with its short form *-ffaa* as it is more economical, frequently occurring ordinal in the texts this study consulted, easy to remember and being short as compared to others.

Apart from the critical analysis of the orthography of Oromo, this study also examined the lexical elaboration of the language from two arenas or spheres (i.e., from the standardization committee and from the primary and secondary textbook writers). In this regard, both the standardization committee of Oromo and textbook writers applied both internal and external means of elaborating the language. Semantic extension, derivation, compounding, blending and abbreviations are among the internal means that are used to elaborate the language in both spheres; i.e. in the documents compiled by the standardization committee of Oromo and in the textbooks. The extent of the use of both internal and external means is not the same in the both spheres.

Regarding, internal means of lexical elaboration, meaning extension, derivation, compounding and abbreviation are the internal means that are extensively used by textbook writers than the standardization committee of Oromo. On the other hand, blending is among the internal means of language enrichments that is equally extensively used by both the standardization committee of Oromo and by the textbook writers.

When we see the elaboration technique a bit detailed, semantic extension is more applied in the textbooks than in the documents of the standardization committee of Oromo. The meanings are extended from parts of a human body parts, animals 'bodies, and foods, farming activities, a plant, materials and shelters. Derivation is also extensively used in both documents to elaborate the lexicon of the language. Nominalization, verbalization and adjectivization are among the strategies of derivation that are applied to elaborate the language. Of these three derivation strategies, nominalization and verbalizations are the most productive derivational means of

elaborating the language (see section 5.2.1.1.3). Abbreviation, acronyms and blending are also other techniques of internal elaboration of the language. Blending is extensively used in the elaboration of the language. Oromo employed blending extensively for the enrichment of the language. On the other hand, as compared to other means, abbreviation and acronyms are less productive in Oromo. The abbreviated forms of two or three words are usually written in Oromo with and without period.

Compounding is also one of the internal means of lexical elaboration which is extensively used in both the documents of the standardization of Oromo and in the textbooks. Of the compounding combinations, noun + noun construction of compound is the most extensively applied; followed by the noun + adjective in both spheres. The combination of adposition + verb and adjective + noun are very rare, and not productive as compared to other word classes. Regarding the constituents of the compound, in both in the documents of the standardization committee of Oromo and in the textbooks compiled by textbook writers involve mainly the categories of nouns, adjectives, adpositional and verbs.

In relation to compound types, nominal, adjectival, adpositional and verbal compounds have been identified in Oromo in the documents consulted in general. Of the three types of compounds, nominal compounds are the most productive; followed by adjectival, adpositional and verbal compounds. Of the nominal compound formation, noun + noun construction is the most productive combinations of the nominal compounds. Adpositional compounds and verbal compounds are less commonly attested in Oromo. Verbal compounds in both the documents of the standardization committee of Oromo and in the textbooks are of composite type (i.e., ideophone + verb 'said /to'), as in for instance, *cal* 'silent' + *jedhe* 'said'. Phonological modification strategies are also employed in between the constituents (in boundaries) of the words.

Concerning semantic heads, compounds of Oromo are categorized into endocentric (in which the head determines the semantics of a compound), appositional (in which both elements contribute to the meaning of a compound) and exocentric compounds (in which the meaning is out of the constituents). The first types of compounds are very common; whereas the second and third

types are very limited in Oromo. Regarding the compound formation in the language, formal changes such as vowel deletion of segments are exhaustively used in Oromo compounds. Following the formal changes like deletion of segments in word boundary and deletion of the terminals of the second constituent and the majority of the compounds are written as single lexical items in Oromo. Compounds in the language are written as both separate and single words. However, there are inconsistencies in writing compounds in Oromo. Compounds are seen to be written with a hyphen between the constituents, with a space between the constituents and as a single word. These inconsistencies should be harmonized in order to minimize variation in the language. Otherwise, these inconsistencies can highly affect standardization of the language.

Apart from the internal means of lexical elaboration we have seen so far, the standardization committee of Oromo and the textbook writers also used external strategies of lexical enrichments extensively. Borrowing and loan translations are among the external means which are extensively applied to enrich the language. These two strategies are more productively used in the textbooks than in the documents of the standardization of Oromo as several scientific and technical terms are designated via these external means. Oromo has borrowed words from various languages through both direct and indirect methods and integrated them into its phonological systems. The loanwords are adapted into the language through both with modification and without modification. Regarding source language preferences, English is the main source of loanwords of Oromo as English is well developed to express scientific and technological concepts followed by Amharic, Arabic, Italian, Swahili and French. Greek loanwords are almost none in Oromo. From Ethiopian languages, Amharic is the only indigenous language used to elaborate Oromo. On the other hand, superfluous borrowings are attested in Oromo despite the existence of native equivalent in the language. It could affect the development and standardization of the language.

When loanwords are integrated into Oromo, non-existent vowels in Oromo are subjected to replacement with the nearest vowels. Likewise, diphthongs which are existent in English loanwords are replaced either by simple vowels or by a glide insertion in the language. Due to the massive borrowing for the designation of scientific, technological, religious and other concepts, some of the non-native consonants are introduced to Oromo. For example, the

phonemes /p/, /v/, /ʒ/, and /z/ were not part of phonological systems of Oromo. But they are incorporated via loanwords. Furthermore, the phoneme /ʒ/ has not yet been represented by any grapheme officially and it is being represented by various graphemes like using <zy>, <zh>, <zn> and sometimes observed to be replaced by the nearest phoneme /dʒ/ i.e., <J>. This may affect standardization of the language.

On the other hand, the word initial consonant cluster is not permissible in Oromo. Hence, such impermissible word initial consonant cluster and more than two consonant clusters in loanwords are adjusted by epenthetic vowel [i]. Vowel insertion, length, addition and replacements are among the phonological adjustments made during the integration of foreign concepts into the language.

Loan translation is also another kind of external source that is extensively used to elaborate the language. The method is extensively applied by the textbook writers and by the standardization committee of Oromo. Loan translation in the language is of two types; literal and non-literal. Literal loan translation is primarily used in Oromo followed by non-literal or conceptual loan translation. Non-literal loan translation can create better understanding of designated expressions for native speakers as concepts are translated into a target language. It is attested that there exists an overlap between the method of loan translation and compounding in Oromo in some of the cases. In Oromo, the word *mana* 'house' is used productively to coin several concepts through loan translation. This productive pattern functions as a loan translation for Oromo. Even though the translated form has similar in structure with Oromo compound words, semantically it can be identified as loan translation.

Finally, this research identified challenges of in standardization process of Oromo. The steady progress of standardization of the language did not normalize it to the extent expected. This is due to the lack of higher level professionals, lower level status of the authority (the status of the standardization committee of Oromo) (Appendix G, FGD1, 61, FGD3, 27 & interview session, 12). Some of what is being codified by the standardization committee of Oromo is not put in practice in the various spheres (in the education sector, particularly in the textbooks). Rather than using the codified forms by the authority uniformly, various individuals and institutions are

establishing their own forms (codes). In this regard, this research identified that there are variations as to how compounds, abbreviations, ordinals, word divisions should be written (see section 5.2.1.2). There are no set rules to guide the variation. Moreover, there is lexical variation that resulted from dialectal variation which has to be regularized (see section 5.2.1.2, appendix F). There is also challenge of reforming the orthography of the language. There is also a challenge of marking of the sounds represented by digraphs when they are geminated. The grapheme representations of the ejectives are also another challenge of normalizing of the language as it has a negative pedagogical impact on the transfer of reading skills (Stark 2010:162).

6.3 Recommendation

Based on the challenges identified in the study on orthographic and lexical standardization of Oromo, the following sets of recommendations are forwarded.

- The alphabetic order of the Roman based script (*Qubee*) of Oromo needs reconsideration and hence, to be reordered via placing digraphs immediately after their respective monographs as it is depicted in appendix c. This is due to the fact that the alphabetic order has a pedagogical implication and any alphabet of a language has to be simple to the learner to identify the characters easily (Stark 2010:157).
- A glottal stop is a frequently occurring sound in the texts of Oromo. It is also included as an independent phoneme in the phonemic inventory of the language. Since the glottal stop is now represented by the apostrophe, the standardization committee of Oromo has to design a grapheme for the sound. This study recommends the IPA glottal symbol <ʔ> for the graphematic representation of the glottal stop and its place in the alphabetic order should be at the end of the alphabet (see Appendix C).
- The standardization committee of Oromo has to reconsider the unsettled issues of the under representation of digraphs in their geminated forms of realization since the issue of the representation of gemination has been one of the factors for the decision to adopt the Roman based script (*Qubee*) to write Oromo. The geminated realization can be marked in two ways; doubling the entire digraph as in the case of the current marking of digraph in

Hadiyya which is not economical (Samuel 2017:103). Another option could be geminating the first constituent of the digraphs. This research recommends the doubling of the first constituent of each digraph as it is more economical as compared to the first option (see section 3.7.3).

- The representation of the ejectives <x> [tʰ], <q> [kʰ], <c> [tʃʰ] and <ph> [pʰ] may have a negative impact on the transfer of reading skills. Therefore, in order to maximize transfer of skills, this research recommends the replacement of the graphemes of Oromo ejectives <x>, <q>, <c> and <ph> by the <tʰ>, <kʰ>, <cʰ>, and <pʰ>, respectively. His study recommends doubling only the first constituent. This is much better than doubling the entire constituent. The other two sounds (i.e., [tʃ] <ch> and [ʃ] <sh>), however, could better be left to be represented using the current working graphemes since there are no challenges of transfer of skills with such graphemes (see section 3.7.4).
- There are no officially codified rules as to how Oromo ordinals, compounds, abbreviations, word divisions should be written. Hence, this research recommends that the standardization committee of Oromo should codify rules to resolve the existing inconsistencies in various texts of Oromo.
- Word break issues are not yet exhaustively codified (normalized) by the standardization committee of Oromo. For instance, there is no convention as to how elements like *wal*, *haa*, *ni*, *kan*, *nan* should be written. These elements are not seen to be consistently written. They are observed to be written sometimes attached to words that appear before or after them. In other instances they are being written as independent words. The present study proposes these elements to be written attached to the words that appear before or after them as the elements do not have meanings independently. Hence, like any other functional elements in the language they better written together with the preceding or words after them. This study also recommends further detailed study has to be conducted so as to set rules regarding word boundary demarcation.

- The current status of the standardization committee of Oromo, most of whose members are having an educational level of masters and first degree and having almost below 13 members, do not much with the existing several tasks of normalizing the language. Hence, this research recommends scaling the current status of authority to the next higher level (academy). The professional with educational rank of professor, doctors has to be the members of the staff so as to truly standardize the language (Appendix G, Interview, 12)
- Some of what is being codified by the standardization committee of Oromo is not put in practice in the various spheres (in the education sector, particularly in the textbooks). Language standardization is a process any language requires the coordination of language standardization stakeholders (such as individual professionals, institutions, media, governmental and non-governmental actors) so as to normalize a language. In this regard, the coordination between the standardization committee of Oromo and all other stakeholders is very loose. Hence this research highly recommends that the committee has to take the lion shares to coordinate individual professionals, education experts, linguists, lexicographers, institutions, media and other concerned bodies so as to work hand in hand (see Appendix G, FGD3, 27, Interview session, 12).
- The education sector has to publish textbooks and other learning materials using a standard Oromo via working in collaboration with the standardization committee of Oromo. All the textbooks and other learning materials have to be distributed from a single center. These authorities also have to work hand in hand to uniformly apply words of different dialects so as to maintain the stable use of the standard forms. In this regard Ethiopian language policy states that government institutions, educational institutions and the mass media shall have special responsibility to maintain the standards of Ethiopia's writing systems that are currently in use (MoCT 2020:23).

6.4. Potential areas of future research

In order to help the standardization process of Oromo, interested researchers can involve in the following language standardization related areas of study.

- An exhaustive descriptive study about variations of the forms and lexical semantics of dialects of Oromo has to be conducted as it has an implication for the standardization of the language
- Vandebussche (2010:1) states that newspapers, radio and television played a great role in the formation, diffusion and acceptance of norms for written and spoken standard Dutch in the 20th century. The author further states that the real massive spread of uniform standard Dutch among the greater part of the population of Flanders only began in the period between the two World Wars. As such, the Flemish population continued to feel the need for normative guidance in linguistic matters throughout the greater part of the 20th century. From Vandebussche's view, we can note that media can play a paramount role in the standardization process of a language. Hence, the role of the media in the standardization process of Oromo is left untouched for upcoming researcher.
- Since the standardization committee of Oromo follows the composite (the inclusion of all dialects) approach in the standardization process of the language, interested researchers can conduct on the perception of users of how much their dialects are included in the standardization process of the language. Apart from the claim of the standardization committee of Oromo that composite (polycentric) approach is followed while normalizing, objective study is vital to validate the claim of the authority. Hence, upcoming researchers can also conduct on the selection process of the base dialect and the fairness of the inclusion of the varieties of the language.
- Knowing the current standardization status of Oromo can play a role to clearly identify what to be done in the future. Therefore, an indepth study on the overall current status of the standardization of Oromo is crucial.
- Finally, since the lexical standardization of a language plays a paramount role, a detailed study about the strategies and inclusion mechanisms of scientific and technical terms have to be conducted. Hence, interested researchers can conduct on the issue.

References

- Abdulsamad Muhammad. 1994. *Seerlugaa Afaan Oromo* [Oromo grammar]. Addis Ababa: Caffee Oromiyaa [tʃæffe Oromia].
- Abu Guba, Mohammed Nour. 2016. *Phonological Adaptation of English Loanwords in Ammani Arabic*. Salford: University of Salford (PhD dissertation).
- Adegbija, Efurosibina. 2004. Language Policy and Planning in Nigeria. *Current Issues in Language Planning* 5 (3). 181-246.
- Adrey, Jean-Bernard. 2009. *Discourse and Struggle in Minority Language Policy Formation: Corsican Language Policy in the EU Context of Governance*. London: Palgrave Macmillan.
- Ager, Dennis. 1996. *Language Policy in Britain and France*. London: Cassel.
- Ager, Dennis. 2001. *Motivation in Language Planning and Language Policy*. Clevedon: Multilingual Matters.
- Ahn, Suzy. 2014. Faithfulness Conflict in Korean Blends. *Proceedings of the 37th Annual Penn Linguistics Conference. University of Pennsylvania* 20 (1). 20-30
- Akanlig-Pare, George & Asante, Rogers. 2016. *Vowel harmony in Nkami*.
<https://www.researchgate.net/publication/303407283>
- Akmajian, Adrian, Demeers, Richard, Farmer, Ann & Harnish, Robert. 2010. *Linguistics: An Introduction to language and communication* (6th edn.). Cambridge: The MIT Press.
- Ali, Mohammed & Zaborski Andrzej. 1990. *Handbook of the Oromo language*. Wroclaw: Polska Akademia Nauk.
- Almaz Wasse. 2018. *The challenges of language standardization: the case of Gamo*. Addis Ababa: Addis Ababa University (PhD Dissertation).
- Amanuel Raga & Samuel Adola. 2012. Homonymy as a barrier to mutual intelligibility among speakers of various dialects of Afan Oromo. *Journal of Language and Culture* 3 (2). 32-43
- Antia, Basse. 2000. *Terminology and Language Planning. An Alternative Framework of Practice and Discourse*. Amsterdam: John Benjamins.
- Appel, Rene & Muysken, Pieter. 1987. *Language contact and bilingualism*. London: Edward Arnold.

- Ariel, Mira. 2002. The demise of a unique concept of literal meaning. *Journal of Pragmatics* 34. 361–402.
- Avanzati, Beatriz & Gutman, Alejandro. 2013. *Description of Oromo*
<http://www.languagesgulper.com/eng/Oromo.html>, (accessed May 21 2018).
- Bahumaid, Showqi. 2015. Lexical Borrowing: The Case of English Loanwords in Hadhrami Arabic. *International Journal of Language and Linguistics*, 2 (6).13-24.
- Baldauf, Richard & Ingram, David. 2001. Language Planning and Language-in-Education Policy. In William Frawley (edr.), *Oxford International Encyclopedia of Linguistics*, 412-416. Oxford: Oxford University Press,
- Baldwin, Timothy & Su Nam Kim. 2010, Multiword expressions. In Nitin Indurkha Fred j. Damerau (edrs), *Handbook of Natural Language Processing*, 2nd edition. 267-292. CRC Press, Taylor and Francis Group, Boca Raton, FL
- Ball, Martin. 2005. *Clinical Sociolinguistics*. London: Blackwell.
- Bamgbose, Ayo. 1989. Issues for a model of language planning. *Language Problems and Language Planning*, 13 (1): 24 – 34.
- Bamgbose, Ayo. 1991. Language and Nation. *Language question in Sub-Saharan Africa paperback*. London: Edinburgh University Press
- Banti, Giorgio. 2008. Oromiffaa. In Siegbert Uhlig (edr.), *Encyclopaedia Aethiopica* (4), Wiesbaden: Otto Harrassowitz.
- Banti, Giorgio & Abdirachid, M. Ismail. 2015. Some issues in Somali orthography. In Cabdirashiid M. Ismaaciil, Cabdalla C. Mansuur and Saynab A. Sharq (edrs) *Somali Language and Literature Studies series* (1). 36-48. Djibouti. The Intergovernmental Academy of Somali Language
- Barnwell, Katherine. 2004. *Developing a Writing System*. Gomba: Seed Company
- Bartsch, Renate. 1985. The influence of language standardization on linguistic norms. *Studia Linguistica* 39: 23–50.
- Bat-El, Outi. 2006. Blend. In Keith Brown, (Editor-in-Chief), *Encyclopedia of Language & Linguistics, Second Edition* (2). 66-70. Oxford: Elsevier.
- Batibo, Herman. 2005. *Language Decline and Death in Africa: Causes, Consequences and Challenges*. Toronto: Multilingual Matters.

- Bauer, Laurie. 2003. *Introducing Linguistic Morphology*, 2nd edn. Edinburgh: Edinburgh University Press.
- Baye Yimam. 1986. *The phrase structure of Ethiopian Oromo*. London: University of London (PhD dissertation).
- Baye Yimam. 1992. ṣirʔatə s'ihifət [Writing systems]. *Dialogue. Journal of the Teachers' of Addis Ababa University* 3,(1). 17–41 (in Amharic).
- Baye Yimam. 1994. Lexical development in Oromo. *Language reform: History and future* 6. 61-72.
- Bekale Seyum. 2012. *Language diversity and challenges of government language planning in Ethiopia*. Addis Ababa: Addis Ababa University (PhD dissertation).
- Benczes, Réka. 2015. Are exocentric compounds really exocentric? *SKASE Journal of Theoretical Linguistics* 12 (3).54–73
- Bender, Lionel. 1972. Loanwords in Amharic daily newspapers. *Anthropological Linguistics* 14 (8). 317-322.
- Bender, Lionel, Mulugeta Etefa & Diane Stinson. 1976. Two Cushitic languages. In Lionel Bender, Donald Bowen, Robert Cooper and Charles Ferguson, (eds), *Language in Ethiopia*, 130-354. London: Oxford University Press.
- Borer, H. 2009. “Hebrew Compounds,” in R. Lieber and P. Stekauer (eds.) *Handbook of Compounds*. Oxford: Oxford University Press.
- Cabré, Teresa. 1999. *Terminology: Theory, methods and applications*. Amsterdam: John Benjamins Publishing Company.
- Cahill, Michael & Elke Karan. 2008. *Factors in designing effective orthographies for unwritten languages*. Texas: SIL internationalelectronic working papers.
- Cameron, Deborah. 1995. *Verbal hygiene*. London. Routledge.
- Central Statistical Authority. 2007. *The Population and Housing Census: Statistical Report for Country level*. Addis Ababa: CSA.
- Cheng, Shiyun. 2018. A Contrastive Analysis of Word Formation of English and Chinese Neologisms. *Theory and Practice in Language Studies* 8, (2). 251-256,
- Clayton, Thomas. 2006. *Language Choice in a Nation under Transition: English Language Spread in Cambodia*. Lexington: University of Kentucky.
- Cobarrubias, Juan. 1983. Ethical issues in status planning. In Juan Cobarrubias & Joshua.

- Fishman (eds.), *Progress in language planning: International perspectives*, 41-85. The Hague: Mouton Publishers
- Cohen, Gideon. 2007. Mother tongue and other tongue in primary education: Can equity be achieved with the use of different languages? In Hywel. Coleman (edr.), *Language and development: Africa and beyond. Proceedings of the 7th International Language and Development Conference Addis Ababa, 26-28 October 2005*, 62-75. Addis Ababa: British Council.
- Cooper, Robert. 1989. *Language Planning and Social Change*. Cambridge: Cambridge University Press.
- Constitution of Oromia regional state. 1995. Finfinnee.
- Coulmas, Florian. 1997. *The Handbook of Sociolinguistics*. Oxford: Blackwell.
- Coulmas, Florian. 1998. Language rights: interests of states, language groups and the individual. *Language Sciences* 20. 63–72.
- Coulmas, Florian. 2007. *Language Governments in Transformation: Future Prospects for German and Japanese in Science, Economy and Politics*. Berlin: Mouton de Gruyter.
- Coulmas, Florian. 2013. *Writing and Society*. Cambridge: Cambridge University Press.
- Crystal, David. 1980. *A first dictionary of linguistics and phonetics*. London: Andre Deutsch.
- Crystal, David. 1984. *Who Cares About English Usage? Harmondsworth*. London: Penguin Books.
- Crystal, David. 1991. *A dictionary of linguistics and phonetics*. Cambridge: Blackwell.
- Crystal, David. 1997. *A dictionary of linguistics and phonetics (6th edn.)*. Oxford: Blackwell.
- Crystal, David. 2003. *English as a Global Language*. Cambridge: Cambridge University Press.
- Decker, Ken. 2014. *Orthography Development for Creole Languages*. [S.l.]: [S.n.]. Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal> (accessed 24 Nov 2019).
- Desalegn Leshyibelu. 2012. *Language and religion: The case of Ethiopian Orthodox Tewahido Church*. Addis Ababa: Addis Ababa University (MA thesis).
- Deumert, Ana. 2003b. Describing language standardization. Models and methods. In K. Arneson (edr.), *Standardisering og variation i vestnordisk*, 9–32). Reykjavik: Iceland University Press.
- Deumert, Ana. 2004. *Language standardization and language change: The dynamics of Cape*

- Dutch. Amsterdam: John Benjamins.
- Deumert, Ana. 2009. Language planning and policy. In Rajend Mesthrie, Joan Swann, Ana Deumert, & William Leap (eds.), *Introducing sociolinguistics*, 371-406. Edinburgh: Edinburgh University Press.
- Education Bureau of Oromia. 2012. *Barnoota saayinsii naannoo (kuta 1-4ffaa)* [Environmental Science textbook (grade 1-4)]. Addis Ababa: Education Bureau of Oromia.
- Edwards, Viv. 2004. *Multilingualism in the English-speaking World: Pedigree of Nations*. London: Blackwell.
- Fasold, Ralph. 1984. *The Sociolinguistics of Society*. New York: Basil Blackwell.
- FDRE. 1995. *Constitution of the Federal Democratic Republic of Ethiopia*. Addis Ababa: Berhanena Selam Printing Press.
- Feda Negesse. 2015. Classification of Oromo Dialects: A Computational Approach. *International Journal of Computational Linguistics (IJCL)*, 6 (1). 1-10
- Felber, Helmut. 1984. *Terminology manual*. Paris: Unesco and Infoterm.
- Ferguson, Charles. 1996. Language development. In Thome Huebner (edr.), *Sociolinguistic perspectives: Papers on language in society*, 40-47. New York: Oxford University press.
- Ferguson, Gibson. 2006. *Language Planning and Education*. Edinburgh: Edinburgh University Press.
- Feyisa Demie. 1996. Historical challenges in the development of the Oromo Language and some agenda for future research. *Journal of Oromo Studies*, 3(1 & 2). 18-27.
- Fikadu Belda. 2010. *Analysis of spelling and grammar errors of teaching materials prepared for Afan Oromo students in Haramaya University*. Addis Ababa: Addis Ababa University (MA thesis).
- Fishman, Joshua. 1970. *Sociolinguistics: A Brief Introduction*. Rowley: Newbury House.
- Fishman, Joshua. 1973. Language Modernization and Planning in Comparison with Other Types of National Modernization and Planning. *Language in Society* 2(1). 23-43.
- Fishman, Joshua. 1975. *Language and Nationalism*. Massachusetts: New Bury House Publishers.
- Fishman, Joshua. 1989. *Language and Ethnicity in Minority Sociolinguistic Perspective*. Clevedon: Multilingual Matters.
- Gadaa Melbaa. 1988. *Oromiya: An introduction to the Oromo People*. Minneapolis: Krik House

Publishers.

- Geeraerts, Dirk. 2003. Cultural models of linguistic standardization. In René Dirven, Roslyn Giegerich, Heinz. 2005. Lexicalism and modular overlap in English. *SKASE Journal of Theoretical Linguistics* 2: 43–62.
- Goldberg, Adele 1995. *Constructions. A construction grammar approach to argument structures*. Chicago: University of Chicago Press.
- Goldberg, Adele 2006. *Constructions at work*. Oxford: Oxford University Press. In Frank & Martin Pütz (eds.), *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*, 25-68. Berlin: Mouton de Gruyter.
- Getachew Anteneh & Derib Ado. 2006. Language policy in Ethiopia: history and current trends. *Ethiopian Journal of Education and Sciences* 2(1). 37-62.
- Gottlieb, Nanette. 2005. *Language and Society in Japan*. Cambridge: Cambridge University press.
- Gragg, Gene. 1976. Oromo of Wallaga. In Lionel Bender (edr.), *The Non-Semitic Languages of Ethiopia*, 166-195. Michigan: African Studies Center.
- Gragg, Gene. 1982. Oromo Dictionary, Bender et al. (eds.), and East Lansing: *International Journal of Ethiopian and Eritrean Studies* 11.223-234.
- Gralow, Francis. 1981. Some sociolinguistic considerations in orthography design. *Notes on Literacy* 33. 8–14.
- Griefenow-Mewis, Catherine & Tamene Bitima. 1994. *Lehrbuch des Oromo*. Köln: Köppe.
- Griefenow-Mewis, Catherine & Tamene Bitima. 1995. *Übungsbuch Oromo*. Köln: Köppe.
- Griefenow-Mewis, Catherine. 1997. *Swahili loan words in Oromo*.
<https://ul.qucosa.de/api/qucosa%3A11711/attachment/ATT-0/> (accessed 16 Sep 2019)
- Griefenow-Mewis, Catherine. 2001. A grammatical sketch of written Oromo. Köln: Köppe.
- Gries, Stefan. 2004. Shouldn't it be breakfunch? A quantative analysis of blend structure in English. *Linguistics* 42 (3). 639-667.
- Guluma Gameda. 2018. *The History and Politics of the Qubee Alphabet*.
<https://k ichuu.com/history- politics-qubee-alphabet/> (accessed 27 Nov 2019)
- Haspelmath, Martin. 2009. *Lexical borrowing: Concepts and issues*.
<https://www.researchgate.net/publication/279973916> (accessed 10 Feb 2020).

- Haugen, Einar. 1959. Planning for a Standard Language in Modern Norway. *Anthropological Linguistics* 1 (3).8-21.
- Haugen, Einar. 1966a. Dialect, language, nation. *American Anthropologist* 68 (4). 922-935.
- Haugen, Einar. 1966b. *Language conflict and language planning: The case of modern Norwegian*. Cambridge, Mass: Harvard University Press.
- Haugen, Einar. 1972. Dialect, language, nation. In Janet Pride & Janet Holmes (eds.), *Sociolinguistics*, 97-11. Harmondsworth: Penguin, 97–111. Originally published in *American Anthropologist* 68 (1966). 922–935.
- Haugen, Einar. 1983. The implementation of corpus planning: Theory and practice. In Juan Cobarrubias & Joshua Fishman (eds.), *Progress in language planning*, 269-289. Berlin:Mouton.
- Hayward, Richard & Mohammed Hassen. 1981. The Oromo Orthography of Shaykh Bakri Sapalo. *Bulletin of the School of Oriental and African Studies* 44. 550-566.
- Heine, Bernd & Kuteva, Tania. 2006. *The Changing Languages of Europe*. New York: Oxford University Press.
- Herbert, Robert.1992. Language in a divided society. In Robert Herbert (edr). *Language and Society in Africa: The theory and practice of sociolinguistics*, 1-19. Johannesburg: Witwatersrand University Press,.
- Hoffer, Bates 2002. *Language Borrowing and Language Diffusion: an Overview*. *Intercultural Communication Studies*.www.academia.edu › Intercultural Communication Studies (accessed 17 Dec. 2019).
- Hoffer, Bates. 2005. Language borrowing and the indices of adaptability and receptivity. *Intercultural communication studies* 14(2). 53-72
- Hornberger, Nancy. 2006. Frameworks and models in language policy and planning. In Thomas Ricento (edr.), *An Introduction to language policy: Theory and method*, 24-41. New York city: Blackwell publishing.
- Hosseinzadeh, Naghmeh Mirzaie. 2014. New Blends in English Language. *International Journal of English Language and Linguistics Research* 2(2). 15-26.
- Hualde, José Ignacio & Zuazo, Koldo. 2007. The standardization of the Basque language. *Language Problems and Language Planning* 31. 143-168
- Hudson, Arnold & Walker, Craven. 1922. *An Elementary and practical Grammar of the Galla*

- or Oromo Language*: London: Society for promoting Christian Knowledge.
- Hudson, Richard. 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.
- International Organization for Standardization (ISO) 704. 1999. *Terminology work-Principles and methods*. International Organization for Standardization.
[www.http.semanticweb.kaist.ac.kr > org > pdocument > standards](http://www.semanticweb.kaist.ac.kr/org/pdocument/standards) (accessed 18 Nov. 2018).
- Jernudd, Bjorn. 1971. 'Notes on Economic Analysis for Solving Language problems'. In Joan Rubin & Bjorn Jernudd, (eds.) *Can Language Be Planned?* 249-263. Honolulu: The University Press of Hawaii.
- Joseph, John. 1984. The elaboration of an emerging standard. *Language Sciences*, 6. 39–52.
- Joseph, John. 1987. *Eloquence and Power: The Rise of Language Standards and Standard Languages*. London: Frances Pinter.
- Kamwangamalu, Nkonko. 1997. Multilingualism and Education Policy in Post-Apartheid South Africa. *Language Problems and Language Planning* 21 (3). 234-253.
- Kaplan, Robert & Baldauf, Richard. 1997. *Language planning from practice to theory*. Clevedon: Multilingual Matters.
- Kaplan, Robert. 2013. Language planning. *Applied Research on English Language* 2(1). 1-12.
- Kebede Hordofa. 1997. Raayaa Oromo Phonology: Aspects of Palatalization. *Ethiopia in Broader Perspectives* 1.469-491.
- Kebede Hordofa. 2005. The Varieties of Oromo. *Journal of Ethiopian languages research center (ELRC)* 1(1). 134-150.
- Kebede Hordofa. 2007. Causative Verb and Palatalization in Oromo. *The Journal of Oromo Studies* 14.95-109.
- Kebede Hordofa. 2009. *Towards genetic classification of the Afan Oromo dialects*. Oslo: The University of Oslo (PhD Dissertation).
- Kenstowicz, Michael. 1983. Salience and similarity in loanword adaptation: A case study from Fijian. *Language Sciences*, 29 (2). 316-340.
- Kessler, Brett & Rebecca Treiman. 2015. *Writing Systems: Their Properties and Implications for Reading*. <https://www.researchgate.net/publication/267393273> (accessed 15 Jan 2020).
- Kloss, Heinz. 1986. *Types of Multilingual Communities: A Discussion of Ten Variables*.

- Sociological Inquiry* 36 (2). 135–145.
- Krapf, Johann. 1840. *Outline of the Elements of the Galla Language*. London: Church Missionary Society
- Krapf, Johann. 1842. *Vocabulary of the Galla[Oromo] Language*. London: Church Missionary Society.
- Lane, Pia. 2018. Language Standardisation as Frozen Mediated Actions: The Materiality of Language Standardisation. In Pia Lane, James Costa, and Haley De Korne (eds.), *Standardizing Minority Languages*, 101-117. London: Routledge.
- Launhardt, Johannes. 1973. *Guide to learning the Oromo (Galla) language*. Addis Ababa: Central Printing Press.
- Leslau, Wolf . 1959. An Analysis of the Harari Vocabulary. *Annales d'Éthiopie* 3. 275-298.
- Lieber, Rochelle, Pavol Štekauer. 2009. Introduction: Status and definition of compounding. In Lieber, Rochelle Pavol Štekauer, eds, *The Oxford Handbook of Compounding* 3-18. Oxford: Oxford University Press.
- Lippi-Green, Rosina. 1997. *English with an accent: Language, ideology & discrimination in United states*. London: Routledge.
- Lloret, Maria-Rosa. 1987. The Morphology of the Causative in Oromo. *Journal of African Languages and Linguistics* 9: 141-156.
- Lodge, Anthony. 1993. *French. From dialect to standard*. London: Routledge.
- Lomaka, Anzhelika. 2017. Literal approach to translation: a classification and literature review. *XLinguae* 10(4). 234-244
- Mabule, Dorah. 2011. *Language Attitudes and Language Choice within the Correctional Services with Reference to Pretoria Central Prison*. Pretoria: University of South Africa (PhD dissertation).
- Magwa, Wiseman. 2008. *Planning for the future: Exploring possibilities of using indigenous African languages as languages of instruction in education – The Zimbabwean experience*. Cape Town: University of South Africa (PhD dissertation).
- Magwa, Wiseman. 2010. Language harmonization in South Africa: Toward a standard unified Shona Orthography for Botswana, Mozambique and Zimbabwe, Midlands State University. *The Dke* (3 (2)). 58-72.
- Mahdi, Hamid Muudee. 1995. *Oromo Dictionary, English-Oromo*. Georgia: Sagalee Oromo

- Publishing Company.
- Malimabe, Refilwe. 1990. *The influence of non-standard varieties on the standard Setswana of high school pupils*. Johannesburg: Rand Afrikaans University (MA thesis).
- Malone, Susan. 2004. *Manual for developing literacy and adult education programs in minority Languagecommunities*. Bangkok: UNESCO.
- Manfred Sailer & Stella Markantonatou. 2018. Multiword expressions: Insights from a multi-lingual perspective (Phraseology and Multiword Expressions 1). Berlin: Language Science Press.
- Masin, Francedsca. 2009. Phrasal lexemes, compounds and phrases: A constructionist perspective. *WORD structure* 2(2). 254-271.
- Mekonnen Hundie. 2002. *Lexical standardization in Oromo*. Addis Ababa: Addis Ababa University (MA thesis).
- Mekuria Bulcha. 1993. Language, Ethnic Identity, and Nationalism in Ethiopia. *The Oromo Commentary* 3 (2). 5-18.
- Mekuria Bulcha. 1994. Priests, Religion and Language in Ethiopia: A Comentary. *The Oromo Comentary* 4 (1). 8-13.
- Milroy, James and Lesley Milroy. 1985. *Authority in Language: Investigating Standard English*. London: Routledge.
- Milroy, James and Milroy, Lesley. 1999. *Authority in Language: Investigating Standard English (3rd edition)*. London: Routledge & Kegan Paul.
- Milroy, James. 2001. Language ideologies and the consequences of standardization. *Journal of Sociolinguistics* 5(4). 530-550.
- Ministry of Youth, Sports and Culture. 1995. *E.C. Cultural policy of Ethiopia*. Addis Ababa: Ministry of Youth, Sports and Culture
- Mirabela, Pop Anamaria & Sim Monica, Ariana. 2009. The use of acronyms and initialisms in Business English. *Journal of English Language (JEL)*. 557-562.
- Mitib Altakhaineh, Abdel Rahman. 2016. What is a compound? The main criteria for compoundhood. *ExELL (Explorations in English Language and Linguistics)* 4 (1). 58-86.
- MoCT. 2020. *Language policy of federal democratic republic of Ethiopian*. Addis Ababa: MoCT.

- MoE. 2012. *Barnoota lammummaa fi amala gaarii kutaa 5ffaa-8ffaa [Civis and Ethical Education (grade 5-8)]*. Addis Ababa: Ministry of Education.
- MoE. 2012. *Keemistirii (kitaaba barataa kutaa 7ffaafi 8ffaa) [chemistry Textbook (grade 7 &8)]* Addis Ababa: Ministry of Education.
- MoE. 2012. *Baayooloojii (kitaaba barataa kutaa 7ffaafi 8ffaa) [Biology Textbook (grade 7 & 8)]* Addis Ababa: Ministry of Education.
- MoE. 2012. *Barnoota Hawaasaa (kitaaba barataa 5ffaa-8ffaa) [Social Study Textbook (grade 5-8)]*. New Delhi: UBS Publishers, Distributors Pvt. Ltd.
- MoE. 2012. *Barnoota Herregaa (kitaaba barataa kutaa 1ffaa-8ffaa [Mathematics students' text book grade 1-8]*. Addis Ababa: Aster Nega publishing enterprise
- MoE. 2012. *Baayooloojii (kitaaba barataa kutaa 8ffaa) [Biology Textbook (grade 8)]*. Addis Ababa: Ministry of Education.
- MoE. 2012. *Afaan Oromoo (kitaaba barataa kutaa 1ffaa-12ffaa) [Oromo Text book (grade 1-12)]* Ministry of Education.]. Kampala: MK publisher
- MoE. 2012. *Afaan Oromoo (kitaaba barataa kutaa 8ffaa) [Oromo Textbook (grade 8)]*: Addis Ababa: Ministry of Education.
- Mohammed Hassen. 1994. *The Oromo of Ethiopia: A History 1570-1860*. Trenton: The Red Sea Press.
- Mulualem Daba. 2017. The Political and Historical Representation of Statues in three Ethiopian consecutive regimes (1889 - Present). *Social Sciences* 6 (2). 35-44
- Mutasa, Elias. 2004. *People Speak: Language Policy and Language Use in South Africa*. Pretoria: S. G. Publishers.
- Myers-Scotton, Carol. 2006. *Multiple Voices: An Introduction to Bilingualism*. New York: Blackwell.
- Nahir, Moshe. 1984. Language Planning Goals: A Classification. *Language Problems and Language Planning* 8 (3). 294-327.
- Nakin, R.M. 2009. *An examination of language planning and policy in the Eastern Cape with specific reference to Sesotho: A sociolinguistic study*. Eastern Cape: Nelson Mandela Metropolitan University (PhD dissertation).
- Ndimande-Hlongwa, Nobuhle. 2010. *Corpus Planning, with Specific Reference to the Use of*

- Standard isiZulu in Media. *Alternation* 7(1). 207-224.
- Onesimus Nesib.1894. *The Galla Spelling Book*.Moncullo (Eritrea): Swedish Mission Press.
- Onesimus Nesib.1899. *The Holy Bible translated in the Galla language*. St. Chrischona: British and Foreign Bible society.
- Owens, Jonathan. 1985. A grammar of Harar Oromo (Northeastern Ethiopia). Hamburg: Buske
- Paffy, Darren. 2008. *Language, discourse and ideology: The real acadamia Spanola and the standardization of Spanish*. Southampton. University of Southampton (PhDissertation).
- Paradis, Carole & Darlene, LaCharité. 2011. Loanword adaptation: From lessons learned to findings. In John Goldsmith, Jason Riggle and Alan C. (Edrs.), *The handbook of Phonological theory(2nd edition)*, 751-778. Oxford: Wiley-Blackwell.
- Penny, Ralph. 2012. *Introduction: language variation*.
https://assets.cambridge.org/97805217/80452/excerpt/9780521780452_excerpt.(accessed 18 March 2020
- Powlison, Paul. 1968. Bases for formulating an efficient orthography.*The Bible translater*19 (2). 74 -91.
- Proclamation No. 46/2001, Enforcement Proclamation of the Revised Constitution of 2001 of the Oromia National Regional State, Megeleta Oromia, Finfine, July 12th 2000.
- Qajeelcha Qormaata Afaan Oromoo. 1998. *Caasluga afaan Oromoo: Jildii I[Oromo Grammar Volume I]*. Finfinnee: Biiroo Aadaafi Turizimii Oromiaatti [Culture and Tourism Bureau of Oromia].
- Qajeelcha Qormaata Afaan Oromoo. 2008. *Jechoota koree waaltinaatiin moggaafaman (Gulaallaman Jildii 1-5)*. Biiroo Aadaa fi Turizimii Oromia.Finfinnee [Oromo study directives. Codified words by standardization committee of Oromo (edited Volume 1-5)].Finfinnee: Culture and Tourism bureau of Oromia.
- Reagan, Timothy. 2002. *Language, Education and Ideology: Mapping the Linguistic Landscape of U.S. Schools*. Westport: Praeger
- Refat, Nadia.2015. Derivation: An Efficient Way of Enhancing English Vocabulary. *International Journal of Language Education and Applied Linguistics (IJLEAL)* 2. 17-26.
- Regourd, Anne. Guest, Nancy Um. 2017. From mountain to mountain: exchange between Yemen and Ethiopia, Medieval to Modern. *Chroniques du manuscrit au Yémen*

- special issue* (1. 42-69).
- Ricento, Thomas. 2006. *An Introduction to Language Policy: Theory and Method*. Oxford: Blackwell.
- Romaine, Suzanne. 1984. On the problem of syntactic variation and pragmatic meaning in sociolinguistics theory. *Folia Linguistica* 18. 409-439.
- Romaine, Suzanne. 1994. 'The sociolinguistic history of *t/d* deletion'. *Folia Linguistica Historica*, 5 (2). 221–55.
- Rosa, Eliane. 2017. Linguistic variation in English. *Revista de Letras* 19 (25). 35-50.
- Rubin, Joan & Jernudd, Björn. 1971. *Can Language Be Planned?* Honolulu: The University Press of Hawaii.
- Rubin, Joan. 1971. Can Language be Planned? In: Rubin, Joan and Jernudd, Bjorn (edr.), *Sociolinguistic: Theory and Practice for Developing Nations*. The Hague: Mouton, 217-252.
- Rubin, Joan. 1973. *Language Standardization and Education*.
[www.eric.ed.gov › products › research-databases › eric](http://www.eric.ed.gov/fulltext/ED114881.pdf). (accessed 20 Jan 2018).
- Sadembouo, Etienne. 1988. "Constitution and function of a language committee and choice of a reference dialect." In Bergman, Ted (edr), *Proceedings of the round table on assuring the feasibility of standardization within language chains*, 10-30. Nairobi: SIL.
- Saeed, John. 2003. *Semantics (2nd edn.)*. Oxford: Blackwell.
- Sager, Juan & Nkwenti-Azeh, B. 1989. *Terminological problems involved in the process of exchange of new technology between developing and developed countries*. Paris: UNSCO.
- Sager, Juan. 1990. *A practical course in terminology processing*. Amsterdam: John Benjamins
- Samuel Handamo. 2017. *Contrastive Analysis of Lexical Standardization in Amharic and Hadiyya*. Addis Ababa: Addis Ababa University (PhD dissertation).
- Sebba, Mark. 2009. *Spelling and Society: The Culture and Politics of Orthography around the World*. Cambridge: Cambridge University Press.
- Schiffman, Harold. 1996. *Linguistic culture and language policy*. London: Routledge.
- Shimelis Mazengia. 2014. *Nominalization via Verbal Derivation Amharic, Tigrinya and Oromo. (A Contrastive Study)*. Addis Ababa: Addis Ababa University (PhD dissertation).

- Shimelis Mazengia. 2016. Some observations on Hadiyyisa orthography. In Binyam Sisay Mendisu & Janne Bondi Johannessen (eds.) *Multilingual Ethiopia: Linguistic Challenges and Capacity Building Efforts*, OSLa 8 (1) 201–218. Oslo: Print House AS
- Smalley, William. 1963. How shall I write this language? *Orthography studies: Articles on new writing systems*, 31-52. London: United Bible Societies.
- Spencer, Andrew. 1991. *Morphological Theory*. Oxford: Blackwell.
- Spolsky, Bernard. 2004. *Language Policy*. Cambridge: Cambridge University Press.
- Standardization committee of Oromo. 1985-2010 E.C. *Wiirtuu [center] Voumel. 1-11*. Finfine: Culture and Tourism Bureau of Oromia.
- Stark, Janie. 2010. *Kambari: orthography*. Texas: SIL international.
- Stern, Judy & Deni Elliott. 1997. *The Ethics of Scientific Research. A Guidebook for Course Development*. Hanover, New Hampshire: Published by University Press of New England,
- Stroemer, Harry. 1995. *A grammar of Boraana Oromo (Kenya): phonology, morphology, vocabularies*. Köln: Rudiger Koppe.
- Takkele Taddese. 2000. *Ways and principles of developing new words and technical vocabulary: The case of Amharic*. Addis Ababa: Addis Ababa University Press.
- Tamam Ahmed Youssouf. 2015. *Issues in the standardization of the Oromo orthography*. Toronto: York University Major Research Paper.
- Tamam Ahmed Youssouf. 2018. *The Copula in the Oromo language (Oromic)*.
<https://www.researchgate.net/publication/322694696>
- Tamene Bitima. 2000. *A dictionary of Oromo technical terms: Oromo–English*. Köln: Köppe.
- Tauli, Valter. 1974. "Theory of Language Planning". In Fishman, Joshua (edr.) *Advances in Language Planning*. The Hague: Mouton.
- Teferi Degneh. 2015. *The development of Oromo writing system*. Kent. University of Kent (PhD dissertation).
- Tekabe Legesse. 2010. *Analysis of Appropriateness and Usage of Physics Terms: Focus on Oromo. Linguistic Appropriateness, Conceptual Precision and Usage of Terms*. Mauritius: VDM Publishing House Ltd
- Temesgen Negessa. 1993. *Word formation in Oromo*. Addis Ababa. Addis Ababa University

- (MA thesis).
- Tench, Paul. 2007. A New Orthography in Tera. *Journal of the international phonetic association* 37. 1-20.
- Tesfaye Shewaye & Taylor, C.V. 1976. Language curricula. In Lionel Bender, James Bowen, Robert Cooper, & Charles Ferguson (eds.), *Language in Ethiopia*, 371- 399. London: Oxford University Press.
- Teshome Lemu. 2014. *Lexical Use of Afaan Oromoo in the Government Mass Media*. Addis Ababa: Addis Ababa University (MA thesis).
- Thomas, Wayne & Collier, Virginia. 2002. *Summary of findings across all research sites. A national study of school effectiveness for language minority students' long-term achievement*. Final Report: Project 1.1. Santa Cruz, CA: Center for Research on Education, Diversity and Excellence, University of California.
- Tiglu Geza. 2015. *Language planning and policy in Silt'e Zone*. Addis Ababa: Addis Ababa University (PhD Dissertation).
- Tilahun Gamta. 1993. Qube Afaan Oromo: Reasons for choosing the Latin script for developing an Oromo alphabet. *Journal of Oromo Studies*, 1(1). 36-41.
- Tilahun Sineshaw. 1997. The Ethiopian Literacy Canon: Ambitions and Frustrations. In: XIII *Proceedings of Ethiopian Studies*, 3. 506-521.
- Tolemariam Fufa. 2009. *A Typology of Verbal Derivation in Ethiopian Afro-Asiatic Languages*. Utrecht: Landelijke Onderzoekschool Taalwetenschap (LOT).
- Tollefson, James. 1991. *Planning language, planning inequality: Language policy in the community*. London: Longman.
- Tollefson, James. 2002. *Language policies in education: Critical issues*. Mahwah, NJ: Lawrence Erlbaum.
- Trudgill, Peter and Chambers, Jack. 1980. *Dialectology*. Cambridge: Cambridge University Press.
- Trudgill, Peter. 1999. Standard English: What it isn't. In Bex, Tony et al (eds.) *Standard English*. London and New York: Routledge.
- Tutin, Agnès. 2008. *For an extended definition of lexical collocations*.
<https://www.researchgate.net/publication/32223490> (accessed 17 Dec. 2019)
- Tutschek, Charles. 1845. *A Grammar of the Galla language*. Munich: MD.

- Vandenbussche, Wim. 2010. Standardisation through the media: The case of Dutch in Flanders. In P. Gilles, J. Scharloth and E. Ziegler (eds.) *Variatio delectat. Empirische Evidenzen und theoretische Passungen sprachlicher variation*, (für Klaus J. Mattheier zum 65. Geburtstag) 309–322. Frankfurt: Peter Lang Verlag.
- Van Huyssteen, Linda. 2003. *A practical approach to the standardisation and elaboration of Zulu as a technical language*. Pretoria: University of South Africa (Doctor of Literature and philosophy dissertation).
- Vujich, Daniel. 2013. *Policy and Practice on language of instruction in Ethiopian. International study of childhood poverty*. www.younglives.org.uk (accessed on October 2019)
- Wach, Szymon. 2013. Calquing English terminology into Polish. *Academic journal of modern Philology* 2. 161-169
- Wako Tola. 1981. *The phonology of Mecha Oromo*. Addis Ababa: Addis Abeba University (MA Thesis).
- Wardhaugh, Ronald. 1992. *An Introduction to Sociolinguistics, Second Edition*. Oxford and Cambridge: Blackwell.
- Wardhaugh, Ronald. 2002. *An Introduction to Sociolinguistics*. Malden: Blackwell.
- Wardhaugh, Ronald. 2006. *An Introduction to Sociolinguistics*. Oxford: Blackwell.
- Webb, Vic and Kembo-Sure. 2000. *African Voices: An Introduction to the Languages and Linguistics of Africa*. Cape Town: Oxford University Press.
- Winford, Donald. 2005. Contact-induced changes: Classification and processes. *Diachronica* 22 (2). 373–427.
- Wolff, Ekkehard. 2000. Language and Society. In Heine, Bernd et al. (edr.), *African Languages, An Introduction*. Cambridge: Cambridge University Press.
- Wolfram, Walt. 2006. Variation and language: overview. In: Brown, k. (edr.). *Encyclopedia of languages and linguistics II*, 333-340. Oxford: Elsevier.
- Wondimu Tegegne. 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.
- Wondimu Tegegne. 2016. The development of written Afan Oromo and the appropriateness of Qubee, Latin script, for Afan Oromo writing. *Historical research letter* 28. 8-14
- Wondimu Tegegne. 2019. The Politics of Qubee in Ethiopia: Rationales and Criticisms to the

- Adaption of Qubee as an Official Orthography of Afaan Oromoo. *Gadaa journal* volume 2 (1). 18-38.
- Wright, Sue. 2004. *Language Policy and Language Planning: From Nationalism to Globalization*. New York: Palgrave Macmillan.
- Wurm, Stephen. 1994. *Language Death and Disappearance: Causes and Circumstances*. Oxford: Berg.
- Zelealem Leyew. 2008. Lexical development in Sidaama. *Journal of Education for Development* 3(1). 97-123.
- Zelealem Leyew. 2009. Lexical development in Tigrinya. *Journal of Education for Development* 3(2). 63-92.
- Zelealem Leyew. 2012. The Ethiopian language policy: A historical and typological overview. *Ethiopian Journal of Languages and Literature Volume* 12(2). 1-59.
- Ziering, Patrick&Vander Plas, Lonneke. 2014. What good are 'Nominalkomposita' for 'noun compounds': *Multilingual Extraction and Structure Analysis of Nominal Compositions using Linguistic Restrictors. Proceedings of COLING, the 25th international Conference on Computational Linguistics: Technical Papers*.1047–1058.
- Ziering, Patrick.2018. *Indirect Supervision for the Determination and Structural Analysis of Nominal Compounds*. Stuttgart: Institut für Maschinelle Sprachverarbeitung der Universität Stuttgart (PhD dissertation).

APPENDICES

Appendix A. FGDs guiding and interview questions

I) FGD1: Guiding questions for members of the standardization committee of Oromo

1. What is your role in the standardization committee of Oromo?
2. Have you taken any training that is related to lexicography and language standardization before engaging into lexicography and standardization of Oromo?
3. Explain if there are orthographic reforms made by the committee or by other concerned bodies? What was the reaction to the reform?
4. Would you please discuss if there are any orthographic representations related problems in Oromo?
5. Which variety(s) is (are) being used as a base dialect in codification processes of Oromo?
6. How do you coin new words for new concepts from the language? Do you have ways to change the codified words whenever you get better words for the concept?
7. Would you please discuss the lexical elaboration techniques you frequently use to designate foreign concepts in to Oromo?
8. If lexical borrowing is the final option, which language(s) do you prefer and why?
9. How do you disseminate the newly coined lexicons to the users?
10. Have you conducted any training or discussion about standardization of Oromo for stakeholders like education and media expertise?
11. Would you explain if there are orthographic related problems being observed in the broadcast media and in the primary and secondary school text books? How do you solve?
12. What do you suggest for further standardization of Oromo?

(Oromo version of FGD 1)

1. Koree waaltina Afaan Oromoo keessatti ga'een kee maali?
2. Hojii Afaan Oromoo waalteessuu keessatti otuu hin hirmaatin duraa leenjii afaan waalteessun walqabtee fudhattanii beektuu?
3. Qubeen Afaan Oromoo waaltina Afaan Oromootin yokaan qaama dhimmii ilaallatun fooya'ee beekaa? Kana dura fayya'ee beeka tanaan yaanni ummataa maal fakkaata ture?

4. Rakkinni haala qubeefi qubeessuu Afaan Oromoo faana walqabtee yoo jiraate otoo ibsitani?
5. Adeemsa waltessuu Afaan Oromoo keessatti looga (loogoota) kam bu'ureffattan?
6. Yaada haaraatif Afaanuma Oromoo keessaa akkamitti jechoota moggaafu? Yaada kana dura moggaafan tokkoof jechi caalatti ibsu yoo argattan jijjirtani beektu?
7. Jechoota afaan alaarraa gara Afaan Oromootittii toftaawaan itti madaksitan otoo ibsitan?
8. Jechoota afaan biraa irraa ergiffachuun dhirqama yoo ta'ee, jechoota afaan kam irra caalaa filattu? Maalif?
9. Jechoota moggaafan akkamitti ummata biraan geessuu?
10. Afaan Oromo waalteesuun walqabatee haga ammatti qaamoota dhimmii ilaallatuf (fakkeenyaaf, ogeeyyi barnootaa fi ogeeyyi subqunnamtiif leenjii kennitanii yokaan waliin marii gootanii beektuu?
11. Haala barreeffama Afaan Oromoon walqabte rakkooleen kitaaba barataa sadarkaa tokkooffaafi midiyaalee biroodkaastiwaan Afaan Oromoo keessatti waan argitan jiraa? Yoo jiraatee, garaagarumma jiruu furuf maaltu hojjatamaa jira?
12. Afaan Oromoo akka gaaritti waaltessuf hojiin kana caalaa hojjatamuu qaba jittan otoo ertanii?

II) FGD 2: Guiding questions for the member of the standardization committee of Oromo

1. Is there any standardized grammar book of Oromo developed by the committee or by Oromo culture and tourism bureau that is being used as codified norms for the users?
2. Explain if you or any concerned government bodies have ever classified the dialects of Oromo?
3. Which dialect (s) is (are) taken as a base dialect in the grammar book you developed or the grammar book developed by Oromia culture and tourism bureau?
4. Explain if you ever made any training, discussion or symposium regarding standardization of the language with expertise of language; particularly with primary textbook developers and media expertise?
5. Discuss if there are morphological and other related challenges which need to be standardized?

6. What do you suggest for further standardization of Oromo?

(Oromo version)

1. Kitaabni Caasluga Afaan Oromoo waalta'e kan qophaa'eefi akka wabitti tajaajilu kan koreedhan yokaan biiroo aadaa fi turizimiitiin qophaa'e jiraa?
2. Loogni Afaan Oromoo koreedhaan yokaan qaama mottummaa dhimmii ilaallatuun qoratamee qoqqoodamee bekaa?
3. Kitaaba caasluuga isiiniin yokaan biiroo aadaafi turizimitin barreeffamee keessatti looga kam bu'ureffattan?
4. Waaltina Afaan Oromootin walqabatee haayulee kitaaba barrataa sadarkaa tokkoffaa qophassanif akkasuma gaazeexeesiitootaafi qamaa dimmi ilaallatuf waaltina irratti leenjiiwanfi marii waliin gaggeessitan yoo jiraatee otoo ibsitan?
5. Ittii fayyadama xiinlatiin walqabatee rakkooleen waaltinaa mul'atan maal akka fakkaatu otoo ibsitani?
6. Afaan Oromoo kana caalaa waalteessuf waan hojjatamuu qabu otoo ibsitani?

III) FGD3:-Guiding questions to education expertise

1. How do you evaluate the current working Roman based script (Qubee) of Oromo?
2. How is Oromo elaborated in the textbook of primary and secondary textbooks?
3. What are the existing challenges of standardization of Oromo in the primary and secondary school textbooks?
4. How do you evaluate the coordination among stakeholders' to standardize Oromo?
5. How do you included (explained) foreign concepts in the textbooks?
6. What do you think needs to be done to standardize Oromo?

(Oromo version of FGD 3 guiding questions)

1. Qubee Afaan Oromoo amma hojirra jiruu kana akkamittii ilaaltu?
2. Afaan Oromoon kitaaba barataa sadarkaa tokkoffaafi lammaffaa keessatti akkamittii gabbataa jira?
3. Waaltina Afaan Oromoo walqabatee, hudhaawwan kitaaboota sadarkaa tokkooffaa fi lammaffaa kessattii mul'atan maal fakkaatu?
4. Waaltina Afaan Oromootiin walqabatee, hariiroo dhimmamtootaa akkamitti ilaaltu?
5. Yaada Afaan Oromorraa adda ta'aan akkamittii ibsiti (moggaaftuu)?
6. Afaan Oromoo waaltessuuf waaltu hojjatamuu qaba jittu?

IV) Interview questions for head of center for culture of Oromia

1. When was Oromo standardization committee established? What is its role?
2. Explain if the committee has ever made a discussion with language experts, media expertise, authors, speakers at large etc.?
3. How do you distribute the codified norms to the users?
4. Have you ever made any assessments (through oral discussion or written) regarding the reaction of Oromo speakers towards standardization of the language?
5. How do you check whether the codified norms have got acceptance (being implemented) or not by language users?
6. Is there any plan to elevate the language in to the federal official working language? If it is to be elevated to the next higher official status, what extra efforts has to be made in standardizing the language?

(Oromo version of the interview questions)

1. Koree waaltina Afaan Oromoo yoom hundaa'ee? Ga'ee isaa maali?
2. Q: Koreen kun waaltina afaanichaatiin walqabatee haga ammaatti haayyuulee adda addaa (fakkeenyaaf haayyuulee barnootaa, miidiiyaa, barressitoota fi kkf) waliin mari'atee beekaa?
3. Kan haga ammaattii waalteessitan akkamittii ummata biraan geessu?

4. Afaan Oromoo waaltessuun walqabatee yaada ummataa kanaan dura qorattaanii yoo beektan ta'e otoo ibsitanii?
5. Afaan Oromoo haga ammaatti waalteessitan fudhatama argachuufi dhabuu isaa akkamittii mirkaneeffattu?
6. Gama motummaatiin Afaan Oromoo afaan hojii motummaa feedaraalaa godhuuf karoorri jiraa? Yoo karoorri jira ta'e, afaanicha kana caalaa waalteessuf tattaaffin kana caalaa godhamu maali?

Appendix B. The current Roman-based Oromo script (Qubee)

Upper case	lowercase	(IPA)
Transcription		
A	a	/a/
B	b	/b/
C	c	/tʃ/
D	d	/d/
E	e	/e/
F	f	/f/
G	g	/g/
H	h	/h/
I	i	/i/
J	j	/dʒ/
K	k	/k/
L	l	/l/
M	m	/m/
N	n	/n/
O	o	/o/
P	p	/p/
Q	q	/q/
R	r	/r/
S	s	/s/
T	t	/t/
U	u	/u/
V	v	/v/
W	w	/w/
X	x	/tʰ/
Y	y	/j/
Z	z	/z/
Ch	ch	/tʃ/

Dh	dh	/d/
Ny	ny	/ɲ/
Ph	ph	/pʰ/
Sh	sh	/ʃ/

Appendix. C. The recommended (by the present study) order of Oromo alphabet

Upper case	lowercase	(IPA)
Transcription		
A	a	/a/
B	b	/b/
C'	c'	/tʃ̣/
Ch	ch	/tʃ/
D	d	/d/
Dh	dh	/d̥/
E	e	/e/
F	f	/f/
G	g	/g/
H	h	/h/
I	i	/i/
J	j	/dʒ/
K	k	/k/
K'	k'	/k̥/
L	l	/l/
M	m	/m/
N	n	/n/
Ny	ny	/ɲ/
O	o	/o/
P	p	/p/
P'	p'	/p̥/
R	r	/r/
S	s	/s/
Sh	sh	/ʃ/
T	t	/t/
T'	t'	/t̥/
U	u	/u/

V	v	/v/
W	w	/w/
Y	y	/j/
Z	z	/z/
Zh	zh	/ʒ/
?	?	/ʔ/

Appendix D. Spelling variation in Oromo

	Orthographic variation	Sources	Total
1) 'multiatom'	atoombaay'ee	Ch. Gr 7 p. 55	1
	atoom-baay'ee	Ch. Gr. 7 p.60 & p.70	2
	atoom baay'ee	Oro. Gr. 9 p. 108	1
	Total		4
2) 'pronoun'	ba-maqaa	Oro. Gr. 9 (1) p. 108	1
	bamaqaa	Oro. Gr. 9 (1) p 12; Oro. Gr. 10 (14) p. 45, p.50, p. 50, p.51, p. 89 (3 times); Oro. Gr. 11 (8) P. 45, p.58, p. 58, p.59 (4 times), 231; Oro. Gr 12 (17), p.77, p.82 (11 times), p.83(2 times), p.159, and p.163 (2 times)	40
Total			41
3) 'manmade'	nam- tolchee	Bio. Gr. 8 (8) p. 9 (2 times), P. 75 (2 times), p 124 (4 times), Bio. Gr. 7 (6) p.13, p.101 (2 times), p.112, p.130(2), Oro. Gr.11 p. 195	15
	namtolchee	Bio. Gr. 8 (1) p. 124, Bio. Gr. 7 (1) P. 16	2
Total			17
4) 'unicell'	seelqeenxee	Bio. Gr. 8 (2) p. 23 &23	2
	seel-qeenxee	Bio. Gr. 7 (2) p.41, p. 45; Bio. Gr. 8 (45) p. 23(2 times), p.24 (9 times), p. 25 (7 times), p.26 (4 times), p.27 (2 times) p.30, p.32, p.34, p.36, p.38 (2 times) p.39, p.48, p.49, p.51	47
	seel qeenxee	Bio. Gr. 7 (1) p. 36	1
Total			50
5) 'revision'	keessa deebii	Ch. Gr. 8 (22) p.5, p. 5, p.6 (4 times), p. 53, p.53, p. 54, p. 55, p. 56, p. 87, p. 87, p.88, p. 89, p. 90, p.91, p. 183, p. 183, p. 184, p.185, p.186 Ch. Gr. 7 (25) p. 5 (3), p.6, p.6, p.7, p.18, p.18, p.19, p.72, p.72, p.73, p.74, p.110, p.111, p.111, p.112, p.113, p.114, p.141 (3), p.143, p.144, p.145	47
	keessadeebii	Ch Gr. 8 (4), p. 91, p.105, p.105, p.106	4

	keessa-deebii	Ch. Gr.8 p.5 (2 times) , p.6 (3 times), p.7,p.57, p. 107,p. 155, Ch. Gr.7 p. 5 (3 times), p.6 (2 times), p.7,p.51, p.75, p. 115	19
	Total		70
6) 'of fake leg'	<i>miil-soboo</i>	Bio. Gr. 8 (12) P. 26, 27 (7 times), p.28, p.38, p.48, p.51	12
	miil soboo	Bio. Gr. 8 P. 27	1
	Total		13
7) biterm	tarmi lamee	Maths Gr.8 P. 39	1
	tarmi-lamee	Maths Gr.8 (20) P. 25 (2 times) , p.27 (2 times), p.39 (3 times), p.40 (2 times), p.41 (2 times), p. 42, p. 43, (6 times), p.50 (2 times)	20
	tarmilamee	Maths Gr.8, P. 36	1
	Total		22
8) 'living things'	lubbu-qabeeyyii	Bio. Gr. 8 (11 times); P. 24, p.122,p.122, p.123, p.126 (4 times),p.143, p.146, p.147 ; Bio. Gr.7 (45 times); P. 8 (4 times), p.9 (4 times), p.10 (3) p. 11 (2 times), p.18, p.22, p.25, p.29 (2times), p.30, p. 31 (5times), p. 34 (2 times), p. 40 (2 times), p. 41, p.45, p. 140 (2 times, p.141 (3 times), p. 142 (2 times), p.146, p.154 (2 times), p.158, p.159 (3 times)	56
	lubbu qabyee	Bio. Gr.8 (7 times) P. 14, p. 21, p.24, p.122, p.139, p.143, p.144; Bio. Gr.7 (11times) P.11, p.14, p.25, p.34, p.54, p.63, p.117, p.140, p.146 (3 times), p.152	18
	Lubbuqabeeyyii	Bio. Gr. 7 (3 times) p. 8, p.9 &p.159; Bio. Gr.8 P. 147	4
	Total		78
9) 'uppercase'	qub-guddeessaa	Oro. Gr. 10 p. 166; Oro. Gr. 9 P. 179	2
	qubguddeessaa	Oro. Gr.10 (5) p. 45, p. 51 (4 times); p. 52; Oro. Gr. 9 (2 times) p. 186, p. 187	7
	Total		
10) 'triangle'	rog-sadee	Maths Gr. 8 (60 times); p. 14.(2),p. 67 (3times), p.87, p.88,p. 90 (8 times), p.91 (6 times), p. 92, p. 93 (2 yimes), p. 94, p. 96 (2 times), p.97 (2 times), p.98 (2 times), p.99, p. 100, p. 102, p. 104 (9	153

		times), p. 105 (2 times), p.106, 107 (4 times), p.120, p. 175, p.177 (6 times), p.178 (2 times), Maths Gr.6 p.143 (5 times), p.145, p.146 (4 times), p.147 (2 times), p. 150 (2 times), p.151, p.165 Maths Gr. 7 (77 times); p. 5, p.9 (2 times), p.13 (2 times), p.14, p.71 (2 times), p. 106, p. 115 (3 times), p.116, p.117 (3 times), p.118 (4 times), p.119 (5 times), p.120 (18 times), p. 121 (9 times), p.122 (3 times), p. 123 (2 times), p.126 (3 times),p. 127 (4 times), p. 128 (2 times), p.129 (4 times), 130 (2 times), p.131 (2 times), p. 154 (4 times)	
	rog sadee	Maths Gr. 8 (4 times); p. 14, p. 91, p. 177 (2 times); Mths Gr.6 (29 times); p. 143,145,147,149,151 (and other 24 instances Maths Gr. 7 p. 119, p. 128	35
	rogsadee	Maths Gr. 6 (118 times) P. 145, p.146, p.147, p.148 (3 times),p. 149 (3 times), p. 150 (4 times), p. 151 (4 times), 152 (4 times), p.153 (3 times), p. 154 (3 times), p. 155 (5 times), p. 156 (15 times), p. 157, p.159 (3 times), p. 160 (7 times), p. 161 (2 times), p. 162 (4 times), p.163 (6 times), p.164 (2 times), p.165 (3 times), p. 166 (6 times), p.167, p. 168 (16 times), p.173 (4 times), p. 174 (8 times), p. 175 (4 times), 176	118
	Total		306
12) 'average'	giddu-galeessa	Bio. Gr. 7 p.8, Bio. Gr. 8 (11) P. 37, p.41, p.70 (4 times), p. 73 (2 times), p.75, and p.137 (2 times); Oro, Gr. 9 (2 times), p. 6 & 32; Oro. Gr. 10 P iv ; Oro. Gr. 11 (3) p. 6, p.107 & p.110; Oro. Gr. 12 (3 times) p. 3, p.65 & p.73	21
	giddu galeessa	Bio. Gr. 7 (8 times) p. 37, p.38, p. 39, p.49, p. 49, p. 50, p.70 and p. 94; Bio. Gr. 8 (17 times)P.p.14, p.29, p.34, p.59, p.60, p.60, p.61, p.61, p.65, p.68, p.70 (4 times), p.75, p.77 and p.91; Oro.,Gr. 9; P 53; Oro. Gr.p.11 (3) p. 106, p.123 & p.185	38
	Total		59

13) 'theory'	yaadiddama	Wiirtuu Vol.10 2005, p.46	1
	yaad-hiddama	Wiirtuu Vol.10 2005, p.147	1
	Total		2
14.'literature'	og-barruu	Oro. Gr.12 p.174	1
	Ogbarruu	Oro. Gr. 9 (7 times), p. 6, p. 126 (2 times), p.140, p.162, p. 181; Oro. Gr.10 (4 times); P. 6, p. 63, P. 68, P.71; Oro. Gr.11, p. iv, Oro. Gr.12 (5 times); p.3, p.8 p. 11, p.124 (2 times);	17
	og barruu	Wiirtuu Vol.10 2005, p.44	1
	Total		19
15) 'title'	mata duree	Oro.Gr.9 (5 times); p.5, p.8, p.18, p.20, p.145 Oro.Gr.10 (10 times); p.5, p.9, p. 67, p.70, p.72, p.99, p.111, p. 131, p. 178 (2 times) Oro. Gr.11 (10 times); p. 9, p.19, p. 46, p.75, p. 129. p. 130, p.175, p.192, p.227, p.228 Oro. Gr.12 (2 times); p.27, p.135	27
	mata-duree	Oro.Gr.9 (4 times); p. 46, p.67, p. 145 (2 times) Oro.Gr.10 (7 times); p.52, p.61, p.62, p.67, p.82, p. 94 (2 times), p. 96 Oro.Gr.11 (7 times); p.144, p.147 (2), p.149, p.190, p.192, p. 220 Oro.Gr.12 (14 times); p.111, p. 147, p. 148, p.157 (times), p.160, p. 169, p.170, p.174, p.176, p.178 (2 times)	32
	Mataduree	Oro.Gr.9 (9 times), p.18 (3 times), p. 105, p.111, p.127, p.128 (2 times). Oro.Gr.10 (8 times), p.32, p. 112, p. 113, p. 114, p. 115, p. 123, p. 130 (2 times) Oro.Gr.11, p.12 Oro.Gr.12 -	18
	Total		77
16) 'similarity'	Walsimannaa	Oro. Gr. 9 (2 times), p. 113, p.117 Oro. Gr. 12; p. 69	3

	wal-simannaa	Oro. Gr. 11 (4 times); p.79, p. 85, p.86, p.95; Oro. Gr.12 (5 times); p.63, p.64, p.67, p.69 (2 times)	9
	Total		12
17) 'inferiority'	gad-aantummaa	Oro. Gr. 10; p 50; Oro. Gr. 11; P. 71	2
	Gadaantummaa	Oro. Gr. 11; P. 97	1
	Total		3
18) 'mass communication'	sab-qunnamtii	Oro. Gr.11, P. 224 (2 times)	2
	Subqunnamtii	Civic & ethical education Gr.7 (2 times) p.52, P.73	2
	sub qunnamtii	Social study Gr. 8 p51	1
	Total		5

Appendix E: Summary of loanwords in Oromo

1. English loanwords in Oromo

	Word	gloss	Source
1.	ampuulii	'ampoule'	Env. Sc. Gr. 1 p. 61
2.	antiibaayotiksiin	antibiotic	Bio. Gr. 7P.11
3.	aspiriinii	'aspirin'	Bio. Gr. 7 P. 25
4.	awutobusii	'autobus'	Env. Sc. Gr. 1 p. 103
5.	salfaayidii	'sulphate'	Ch. Gr. 7 P. 25
6.	ayireenii	'iron'	Bio. Gr 7 P. 10
7.	ayodiinii	'iodine'	Bio. Gr. 7 P.10
8.	ayyoodayidii	'iodide'	Ch. Gr. 7 P.25
9.	baayoloojii	'biology'	Bio.Gr. 7 p.1
10.	bilookeettii	'block'	Env. Sc Gr.1 p.51
11.	biroominii	'bromine'	Ch. Gr. 7 P. 25
12.	boonii	'botany'	Bio. Gr.7 p.3
13.	chaartii	'chart'	Bio. Gr. 7.P.26
14.	dijitii		Maths Gr. 3 P. 60
15.	dimookraataawaa	'democratic'	Soc.St Gr.8 p.1
16.	disinfeektaantii	'disinfectant'	Bio. Gr. 7 P.25
17.	diyaameetirii	'diameter'	maths Gr.3 p. 122

18.	elektirikii	'electric'	env. sc Gr.1 p.10
19.	fangasii	'fungus'	bio Gr.7 p.7
20.	fedaraalaawaa	'federal'	soc.st. Gr.8 p.1
21.	foormullaa	'formula'	maths Gr. 7 p. 123
22.	firaakshiinii	'fraction'	maths Gr. 8 p. 77
23.	gaazii	'gas'	env. sc. Gr 1 p. 61
24.	giraama	'gram'	maths Gr.3 p. 82
25.	haayidiroojinii	'hydrogen'	ch. Gr. 7 p.25
26.	ikoology	'ecology'	bio. Gr.7 p.3
27.	injiinara	'engineer'	ev. sc Gr.1 p.39
28.	iskaarvii	'scurvy'	bio. Gr. 7 p.10
29.	iskuweerii	'squar'	maths Gr.3 p. 111
30.	istoovii	'stove'	env. sc. Gr.1 p.54
31.	ji'ology	'geology'	bio. Gr.7 p.5
32.	kaameeraa	'camera'	bio. Gr. 7p.19
33.	kaarboohayidiretii	'carbohydrate'	bio. Gr. 7 p.10
34.	kaarboondaayooksaayidii	'carbon dioxide'	ch. Gr. 7 p.25
35.	kaarboonii	'carbon'	ch. Gr. 7 p.25
36.	kiillogiraama	'kilogram'	maths Gr.3 p. 82
37.	kiloorinii	'clorine'	ch. Gr. 7 p. 25
38.	kompaasii	'compass'	maths Gr. 8 p. 85
39.	kompaawundii	'compound'	bio. Gr.7. p.30
40.	koompaas	'compas'	maths Gr. 3 p. 122
41.	ko'oordiineetii	'coordinate'	maths Gr. 8 p.68
42.	koopparii		ch. Gr. 7 p. 25
43.	kuwaadraantii	'quadrant'	maths Gr. 8 p.71
44.	kuwaashorkorii	'kwashiorkor'	bio. Gr. 7 p.10
45.	laatesii(lattice)	'lattice'	maths Gr. 3 p. 90
46.	leensii	'lence'	bio. Gr.7 p.19
47.	maayikiroo-orgaanizimoota	'microorganizes'	bio. Gr 7 p.10
48.	maayikrooskooppii	'microscope'	bio. Gr. 7.p.29
49.	meerkurii,	mercury	bio. Gr 7 p.13
50.	meetira	'metre'	maths Gr.3 p. 82
51.	miliimeetira	'milimetre'	maths Gr.3 p. 82
52.	miliyoona	'million'	soc.st Gr.8 p.5
53.	mobaayiliii	'mobile'	env. sc. Gr. 1 p. 59
54.	moodeela	'model'	env. sc. Gr.1 p.75

55.	nagaatiivii	'negative'	maths Gr.8 p.69
56.	oksaayidii	'oxide'	chem. Gr. 7 p25
57.	ovaarii	'ovary'	bio. Gr. 7 p. 97
58.	paaraalaloogiraamii	'parallelogram'	maths Gr.3 p. 118
59.	paraaleloogiraamii	'parallelogram'	maths Gr.3 p. 115
60.	paraaleloogiramii	'parallelogram'	maths Gr. 8 p. 92
61.	parpandikulaarii	'perpendicular'	maths Gr.3 p. 109
62.	pirobileemota	'problems'	maths Gr. 3 p. 54
63.	pirootiinii	'protine'	bio. Gr. 7 p.10
64.	pirootoonii	'proton'	ch. Gr. 7 p. 87
65.	pirootiraakterii	'protractor'	maths Gr. 8 p. 85
66.	piropporshinaalummaa	'proportionality'	maths Gr. 8 p. 86
67.	pootaasiyemii	'potassium'	ch. Gr. 7 p. 25
68.	poozatiivii	'positive'	maths gr.8 p.69
69.	raadiyesii	'radius'	maths Gr.3 p. 122
70.	raadiyoo	'radio'	env. sc. Gr. 1 p. 59
71.	raashinaalii	rational	maths Gr. 8 p.
72.	radiyaasii	radius	maths Gr. 8 p. 104
73.	reektaangilii	'rectangle'	maths Gr.3.p. 115
74.	reeshoo	'ratio'	maths Gr. 8 p. 86
75.	rikkeetisii	'rickets'	bio. Gr 7 p.10
76.	ripubilikii	republic	soc.st Gr.8 p.1
77.	rohombosii	'rhombus'	maths Gr.3 p. 115
78.	saayinsii	'science'	bio. Gr.7 p.2
79.	saaykiliikii	'cyclic'	maths Gr. 8 p. 116
80.	sabmariinii	'submarine'	bio. Gr. 7.p.20
81.	santiimeetiraa	'centimetre'	maths Gr. 3 p. 109
82.	seeloota	'cells'	bio. Gr. 7 p.82
83.	sektarii	'sector'	maths Gr. 8 p. 106
84.	sheelfii	'shelf'	env. sc. Gr 1 p. 59
85.	siliippeerii	'slippery'	env. sc. Gr.1 p.49
86.	soodiyeem	'sodium'	ch. Gr. 7 p. 25
87.	soodiyemkiloraaayidii	'sodium chloride'	ch. Gr. 7 p.25
88.	taaksonoomii	'taxonomy'	bio. Gr.7 p.5
89.	taanjantummaa	'tangency'	maths Gr. 8 p. 103
90.	teeknooloojii	'technology'	bio.Gr. 7 p.1

91.	televiziyinii	'television'	env. sc. Gr.1 p.18
92.	tiraafikii	'traffic'	env. sc. Gr.1 p.75
93.	tiraappiiziyemii	'trapezium'	maths Gr.3 p. 115
94.	tirakeraan	'tractor'	env. sc. Gr.1 p.38
95.	tishartii	't-shirt'	env. sc. Gr.1 p.47
96.	Tiyooramii	'theorem'	maths Gr. 8 p. 93
97.	vaayitaamin	'vitamine'	bio. Gr. 7 p.10
98.	verteksii	'vertex'	maths Gr. 8 p. 111
99.	yuunitii	'unit'	maths Gr. 8 p. 78
100.	zeebraa	'zebra'	env. sc. Gr 1 p. 74.
101.	zeeroo	'zero'	maths Gr.8 p.69

2. Amharic loanwords in Oromo

	word	gloss	Sources
102.	baaqeelaa < ባቄላ /bak'ela/	'bean'	Oro. gr.3 p. 120
103.	bataskaana < ቤተክርስቲያን/betəkiristiyan/	'church'	Env. sc. Gr. 4 p117
104	bo'oo < ቦይ /boy/	'canal'	Oro. Gr 12 P.99
105	boqoolloo < በቆሎ /bək'k'ollo/	'maize'	Env. Sc. Gr. 4 P33
106	burtukaana < ብርትኳን /birtik ^w uan/	'orange'	Bio.Gr.7 P.20
107	cammaa < ጭማ /ɸ'amma/	'shoe'	Env. Sc. Gr. 4 p. 17
108	cinqii < ጭንቅ /ɸ'ink'/	'distress'	Wiirtuu 2005 p.5
109	daabboo < ደቦ /dabbo/	'bread'	Bio.Gr.7 P.99
110	dabbara < ደብተር /dəbtər/	'exercisebook'	Oro. Gr.8 P.2
111	danbii < ደንብ /dənb/	'regulation'	Civ.& Eth.Edu Gr. 5 p.83
112	daqiiqaa < ደቂቃ /dək'ik'a/	'minute'	Maths Gr.6 P.52
113	dinnichaa < ድንች /dintʃ/	'potato'	Soc. St. Gr.8 P.9
114	diiqaalessuu < ማዳቀል /madak'əl/	'to hybridize'	Bio .Gr. 7 P.9
115	fishkaa < ፊሽካ /fiʃka/	'whistle'	Wiirtuu 2005 p.7
116	gabaa < ገበያ /gəbəya/	'market'	Env. Sc. Gr.4 p.113
117	gibira < ግብር /gibir/	'tax'	Soc. St. Gr. 7 P.32

119	gindillaa	< ጊንዲላ /gindilla/	'stout'	Bio. Gr.7 P. 15
119	habaqii /ibbiqi	< እብቅ /ibk'/	'dirty material in teff'	Env. Sc.Gr.3 P.37
120	idaa	< እዳ /ida/	'debt'	Wirtuu 2005 p 12
121	kaffaluu	< መክፈል /mækfəl/	'to pay'	Oro. Gr 9 p.43
122	kataraa	< ከተራ /kətəra/	'holding water'	Bio. Gr.7 P. 15
123	kitfoo	< ክትፎ /kitfo/	'mincedmeat'	Soc. St Gr. 6. P
124	laastikii	< ላስቲክ /pīlastik/	'plastic'	Wiirtuu 2005 p.14
125	makara	< መክር /mækər/	'harvest'	Wiirtuu 2005 p.16
126	masqala	< መስቀል /mæskəl/	'cross'	Oro. Gr. 12 P 92
127	qaraxa	< ቀረጥ /k'ərət'/	'excise'	Oro. Gr. 11 P.117
128	qumxaa	< ቁምጥ /k'umt'a/	'shorts'	Soc. St.Gr.5 P.61.
129	sanbata	< ሰንበት /sənbət /	'sunday'	Oro. Gr.9 P.23
130	shittoo	< ሽቶ /ʃitto/	'perfume'	Oro. Gr . 10 p. 77
131	shuguxii	< ሽጉጥ /ʃigut'/	'pistole'	Civ. & Eth. Edu Gr. 7. P.40
132	taraa	< ተራ /təra/	'turn'	Wiirtuu 2005 p.5
133	waqtii	< ወቅት /wək'it/	'season'	Bio.Gr. 7 P.14
134	warqii	< ወርቅ /wərk'/	'gold'	Soc.St. Gr.7 P. 15

3. Arabic loanwords in Oromo

	Word	Gloss	Sources
135	adaba	'good manners'	Oro. Gr.11. p .160
136	ajandaa	'agenda'	Soc. St. G 8 p 60
137	alkoolii	'alcohol'	Bio. Gr.7 P.31
138	araqee	'liquor, alcohol'	Bio.Gr.8 P. 38
139	baabura	'train'	Env. Sc. Gr.4 P. 115
140	baadiyaa	'peasant (rular)'	Oro. Gr 12. p. 28
141	darzanii	'dozen'	Ch Gr. 8 p161
142	dinaara	'dinar'	Maths Gr.8 P.133
143	farasullaa	'items of 17 kg'	Soc.St. Gr. 6. p45
144	haqa	'truth'	Oro. Gr.12 p. 10
145	jabanaa	'kettle (coffee pot)'	Env.Sc Gr. 2 p36

146	kaawuyyaa	'ironing press'	Env. Gr. 3 P.37
147	kibriitii	'match'	Env. Sc. Gr. P.36
148	kiisii	'pocket'	Oro.Gr11 p.198
149	kiraa	'rent'	Oro. Gr. 10 p. 159
150	kitaaba	'book'	Oro. Gr. 12 P.14
151	maqasii	'scissors'	Bio. Gr. 7. P.128
152	masgiida	'mosque'	Civ. & Ethic. Edu. Gr. 7 15
153	mismara	'nail'	Ch. Gr. 8 p. 6
154	mizana	'balance (scale)'	Maths Gr.4 P.63
155	muuzii	'banana'	Env. Sc. Gr. P.1
156	nooraa	'limestone'	Ch. Gr. 8. p3.6
157	qalbi	'grace'	Oro. Gr. 12 P.56
158	sa'atii	'watch'	Maths. Gr.2 P. 34
159	saamunaa	'soap'	Bio. Gr. 7 P. 64
160	shaayii	'tea'	Civ. & Ethic. Edu. Gr. 7 p 43
161	shinii	'cup'	Gr 2 En.Sc p36
162	shukkaara	'sugar'	Env. Sc. Gr. 4 p 101
163	suuqiiwwan	'shops'	Oro. Gr. 11 P167
164	waraqaa	'paper'	Env.Sc Gr 1 p14
165	xaawulaa	'plank'	Bio Gr. 7 P128
166	zayitii	'oil'	Bio. Gr. 8 p. 3

4. Swahili loanwords in Oromo

	Oromo	gloss	Sources
167	duka	'shop'	Oro. Gr.6 P.36
168	gaarii	'cart'	Oro. Gr. 11 p 191.
169	jembee	'hoe'	Env. Sc. Gr. 3 p.51
170	kiti	'chair'	Soc. St. Gr. 4 P.57
171	misinga	'millet'	Env. Sc. Gr. 4 p.93
172	munada	'auction'	Civ.& Eth.Edu. Gr.7 p.51
173	sookaa	'axe'	Env. Sc. Gr. 4 p.41

5. Italian loanwords in Oromo

	Word	gloss	Source
174	boombaa	'pipeline'	Oro. Gr. 5. p.67
175	gaazeexaa	'news paper'	Oro. Gr. 11 P. 218
176	gommaa	'tyre'	Soc. St. Gr.5 p.53.
177	kaartaa	'map'	Civ& Eth. Educ. G 8 p 12

178	kaartoonii	'card board'	Bio. Gr. 7 P.128
179	koreentii	'electric'	Env. Sc. Gr.3 P. 25
180	paastaa	'posta'	Env. Sc. Gr.4P.10
181	poostaa	'posta'	Oro. Gr. 10 P.170
182	salaaxaa	'salad'	Env. Sc. Gr. 4 P. 5
183	simintoo	'cement'	Env. Sc. Gr. 4 p115
184	teeknika	technique	Ch Gr. 7 p. 67

6. French loanwords in Oromo

185	baabura	'train'	Env. Sc. Gr. 4 p. 115
186	biiroo	'bureau'	Env. Sc. Gr.4 p. 138
187	foolisii	'police'	Env. Sc. Gr.4 p. 138
188	miiliyoona	'million'	Env. Sc. Gr. 4 p. 90
189	ripaabilikii	'republic'	Bio. Gr. 8 p. 1
190	saantima	'cent'	Maths Gr. 2 P97
191	sijaaraa	'cigarette'	En Sc Gr. 4 p33

7. Greek loanwords in Oromo

	Word	Gloss	Source
192	maaraatoonii	'marathon'	Env. Sc. Gr. P. 127
193	Itoophiyaa	Ethiopia	Env. Sc.Gr. P.169 127

Appendix F Lexical variation due to dialectal variation (Sample)

	western	northern	central	southern	eastern	Southeaster n	Gloss
1	ammo	ammo	immoo	ammoo	immoo	ammo	'also'
2	arfaasaa	Furma	afraasaa	badheessa	badheessa	arfaasa	'spring'
3	<i>cinnaa</i>	<i>maddii</i>	<i>cinnaa</i>	<i>moggaa</i>	<i>moggaa</i>	<i>maddii</i>	'side'
4	dansa	gaarii	gaarii	bayeessaa	misha	misha	'good'
5	daraara	daraaraa	daraaraa	daraaraa	ilillii	daraaraa	'flower'
6	<i>Dilbata</i>	<i>Dilbata</i>	<i>Sanbatagudda</i> <i>a</i>	<i>Dilbata</i>	<i>Dilbata</i>	<i>Dilbata</i>	'Sunday'

7	firoomsee	durdubee	durga/dubga	durdubee	durdubee	firoomsee	'preposition'
8	hojji	dalagaa	hojii	dalagaa	hujii	hojii	'work'
9	ishee	ishii	ishee	ishee	ishee	ishii	'she'
10	jecha tishoo	jecha makaa	jecha mkaa	digala	jacha makaa	jecha tishoo	'compound'
11	Jimaata	Juma'a	Jimaata	Jimaata	Juma'a	Juma'a	'Friday'
12	Kibxata	Kibxata	Lammaffo	Lammaffo	Kibxata	Kibxata	'Tuesday'
13	kijiba	kijiba	soba	kijiba	kijiba	dhara	'false'
14	killee	buphaa	buphaa	killee	hanqaaquuu	hanqaaquuu	'egg'
15	lukkuu	lukku	andaqqoo	hindaqqoo	lukkuu	lukkuu	'hen'
16	mara	hunda	Hunda	cufa	cufa	hunda	'all'
17	nafa	jismii	Dhagna	qaama	nafa	nafa	'body'
18	qarshii	birrii	Birrii	maallaqa	maallaqa	maallaqa	'money'
19	Roobii	Roobii,	Roobii,	Roobii,	Roobii	Yaato	'Wednesday'
20	sanbata	Sanbata duraa	Sanbata	Sanbata xinnaa	Sanbata xinnaa	Sanbata xinnaa	'Saturday'
21	waanjo	jeedallo	Geedala	ukkullo	ukkullo	sardiida	'fox'
22	walgayii	walgayii	walgayii	bayii	yaa'ii	kora	'conference'
23	wayya	huccuu	Kafana	huccu	huccuu	wayya	'cloth'
24	Wiixata	Wiixata	Hojjaduree	dafinoo	Dafinoo	Dafinoo	'Monday'
25	woggaa	umrii	Oggaa	ganna	ammata	ganna	'age'

Appendix G: Transcripts of focus group discussions and interview

I) Transcription of FGD 1

1. Qorataa:- (Q-kana booda) Koree waaltina Afaan Oromo kessatti ga'een kee maali?

Researcher (R- henceforth): What is your role in the standardization committee of Oromo?

2. Hirmaataa 1:-(H- kana booda) Galatoomi. Koree waaltina Afaan Oromoo kessatti haga bara 2014 miseensa ogeessa waaltina Afaan Oromo ta'ee tajaajilan ture. Bara 2015rraa jalqabee ammo dursaa garee moggaasa jechootaafi waaltina Afaan Oromooti.

Participant 1 :(P1 henceforth): Thank you. Till 2014, I have been working as one of the members of expertise of the standardization of Oromo. I am head of the department head of lexical codification and standardization of Oromo from 2015 onwards.

3. H2:-Anii ogeessa afaanifi miseensa koree waaltina afaanichaatii

P2:- I am an expert of the language and a member of the standardization committee of Oromo.

4. H3:- Anii ogeessa afaanifi miseensa koree waaltina afaanichaatii

P3:-I am also language expert and a member of the standardization committee of Oromo.

5. H4: Ani Ogeessa Afaan Oromoofi koree waaltina Afaani keessatti miseensa.

P4: I am an expert of Afaan Oromo and a member of the standardization committee of Oromo.

6. H5:- Anis ogeessa Afaanichaafi miseensa

p5: I am also an expert of the language.

7. H6:- akkuma isaani ogeesuma afaanichaati

P: Me too an expert of the language.

8. Q: - Hojii Afaan Oromo waalteessuu keessatti otuu hin hirmaatin dura leenjii afaan waalteessun walqabtee fudhattanii beektuu?

R:- Have you taken any training that is related to lexicography and language standardization before engaging into lexicography and standardization of Oromo?

9. H1:- Ega yeroo sanitiiani barumsa edilee digrii tokkoffaa enaa baradhu koorsiin waaltinarratti xiyyeeffatu hin jiruu ture. Barnoota afaani yoo baranu irra caala hogbarrurratti xiyyeeffata ture. Digrii koo isa tokkoffaa ogaan baradhu waa'ee waaltina afaani akka korsiiitti ogaan barachuu

baadhelee mataduree adda addaa kessattii waa'ee karoora afaani, waaltina afaani, amaammata afaani, waa'ee moggaasa jechootaa baradhees dubisees tureera. Haa ta'utii garuu digrii lammaffaa yoon baradhuu waa'ee waaltina afaanii, saayinsiisaas maalummaa waaltina afaanii, barbaachisummaa isaas, faayeedaa waaltina afaani sirriiti bara dheera barreesseras. Digrii lammaffaa yoon barachaa turee sana waa'ee koree waaltina Afaan Oromoofi waaltina Afaan Oromoorrattii ' term paper' qopheessen ture.

P1:- Ok. There was no standardization related course while I was attending my undergraduate study. The then courses of language focused on more of literature. But during my post graduate, though language standardization was not given as course, there were sub topics on language standardization and lexical standardization in particular, language planning, and language policy. I have learnt the meaning of language standardization, theory of language standardization, function of language standardization and the need for language standardization. In addition to this, I have also prepared a term paper on the standardization committee of Oromo and standardization of Oromo.

10. H2:- Koorsii keessattii akka mata duree tokkotittan fudhadhee malee afaan waaltessurratti leenjii hinfudhanne.

P2:- I have learnt about language standardization as sub topic of a course; but I have not taken any training regarding language standardization.

11. H3:- Afaan sadarkaa digrii tokkoffaa yoo baradhu waa'ee waaltina afaanii sadarkaa koorsiikeessaatti fudhadheen ture. Deebisee ammo afaan karoorsufi waaltessuukan jedhu baratoota digrii tokkooffaaf kenneen turee; ammas kanaan jira.

P3: When I was attending my undergraduate, I have taken it in a course about language standardization. In addition to this I have taught language standardization and language planning course to undergraduate students and I am still teaching.

12. H4:- Akka koorsitti hinfudhannee. Akka mataduree barumsaatitti baradheera

p4: I haven't taken it as a course, but as a topic.

13.H5:- Ani digrii lammaffaa yeroon baradhuu koorsii keessattii fudhadheera.

P5: I have taken it in a course about language when I was attending my MA degree.

14. H6. Akka koorsitittii hin fudhannee.

p6: I haven't taken it as a course

15. Q: Qubeen Afaan Oromoo waaltina Afaan Oromootiin yokaan qaama dhimmii ilaallatuunfooya'ee beekaa? Kana dura fooya'ee beeka tanaan yaanni ummataa maal fakkaata ture?

R: Explain if there is orthographic reform made by the committee or by other concerned bodies? What was the reaction to the reform?

16. H2: Kana dura koree waaltina Afaan Oromootiin qubeen Afaan Oromoo fooya'ee hinbeeku. Biroon barnoota Oromiyaa garuu otoo nuun hinmari'atiin, waan dhimmii isaan hin ilaallatiin tartiiba qubee Afaan Oromo 'A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, Ch, Dh, Ny, Ph, SH' gara'L, A, G, I, M, S, N, U, T, R, E, F, O, D, H, K, J, X, Dh, Q, Y, W, N, B, Ch, Ny, C, P, Sh, V, Ph, Z' tittijjiirani turan.

P2: The standardization committee of Oromo did not make any reform on the working orthography of Oromo since 1991. It was Oromia education bureau, for which the right and the responsibility was not given by the state government of Oromia that reformed the order of the alphabets of Oromo. They reformed it without consulting us. Before the reform was made, the alphabet order of Oromo was 'A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z, Ch, Dh, Ny, Ph, SH' and they reformed it into 'L, A, G, I, M, S, N, U, T, R, E, F, O, D, H, K, J, X, Dh, Q, Y, W, N, B, Ch, Ny, C, P, Sh, V, Ph, Z'.

17. Q:-Tartiiba qubee ulaaga maaaliin jijjiiran?

R: What was the criterion for changing the order of the alphabet?

18. H6: Tartiiba qubee kana kan jijjiiran koorpasii fudhachudhani. Koorpasii sana keessaa qubee kamtuu irra caalaan jecha ittiin barreeffame jedhaniitii tartiiba qubee kan lafa ka'an. Tartiibuma

kanaan kitaabnii barataa sadarkaa tokkoffaarraa jalqabee haga sadarkaa lamaffaatti ittiin maxxanfame barattooniis haala kanaan waggaa lamaaf ittiin barataniituran.

P6: The reason behind the reform, they claim, is the frequency with which each of these characters are used to write a word in Oromo. First, they form a corpus through collecting various texts of Oromo. Then, the bureau puts the most frequent character first with which words in Oromo are written and the least frequent character last. Textbooks of primary and secondary schools were republished due to this reform.

19. R: Jijjirama tartiba qubeetiin walqabatee yaanni ummataa maal ture?

R: How do the speakers of the language react to the change of the order of the alphabet?

20. H5: Qubeen jijjirame waggaa lama caalaa tajaajilarra hin oolle. Ummanis wa'ee kanaa mormii kaaseeti ture. Tariibni kun ammoo adeemsa baruufi barsiisuuf mijataa waan hintaaneeff, ummataanis fudhatama waan hin arganneeffi qaamni dhimmii ilaalatus waan hinmirkaneeessineeffi sababa kana kana fakkaatun tartiba duraanititii debii'era. Otoo fooya'insii qubee afaanichaa barbaachisees kan qoratee motummaaf dhiyeessee kan fooyyessu qaama kana ture.

P5: But the reform did not last more than two years. There was a grievance from the speaker side. The reformation made by the education bureau of the regional government of Oromia was rejected and returned to the previous one for several reasons. Firstly, the education bureau did not consult the responsible body, standardization committee and other concerned bodies. Secondly, the opinion and acceptance of the speakers were not taken into consideration. Thirdly, the maximum ease of learning and other issues were not consulted and considered. If reformation is needed at all, the right of reforming the orthography in general through conducting a research and presenting it to the highest body of the state government is given to us by the state government of Oromia

21. H1: Qubee moggaasuu faana waalqabatee heerrri biyya keenyaa keewwata shanirrattii lafa ka'ameera. Haaluma kanaan calqaba motummaan naannoo Oromiyaa yeeroo hundaa'e Afaan Oromoon Afaan hojii Oromiyaa akka ta'ufi qubeedhanis akka barreeffamu qabu heera

motummaa nannoo Oromiyaa keewwata shaniratti murteeseera. Motummaan naannoo Oromiyaa Afaan Oromoon qubee kamiin akka barreeffamu otoo hinmurteessiin dura haayyuleen gara garaa dhimma kanarratti marii bal'aa taasisaniru. Afaan Oromoon bara 1991A.L.A dura barreesitoota adda addaatiin qubee Saabaa, qubee Arabaafi Laatiniin barreeffamaa ture. Haayyoonni Oromoo kan akka Onismas, Asteer Gannoo, Mahaammad Rashaad, Haylee Fidhaa, Bakrii Saaphaloofi dhaabileen siyaasaa kan akka Adda Bilisummaa Oromoo (ABO) haala barreeffama Afaan Oromoorrattii tattaaffii godhanii haala barreeffama ammaatiif bu'ura ta'aniru. Haaluma kanaan haayyuleen biyya keessaa, biyyaa alaafi dhaabbilee barnoota adda addaarra walumaagalattii qaama dhimmii ilaallatuu hundii walittii dhufaani Afaan Oromoon qubee kamiin otoo barreeffame afaanichaaf mijata, guddisas jedhani marii bal'aafi qorannoo adda addaa ega gaggeessani booda qorannoo sana motummaa naannoo Oromiyaaf dhiyeessanii turan. Qorannoon sun motummaaf yeroo dhiyaatuu yaada adda addaatu dhiyaatee.

22.H3:-Qubee Arabaan, Saabaatiinfi Romaatiin haabarreeffamu kan jedhu yaadni adda addaa ka'eet ture. Ega mariin bal'aan taasifamee booda qubeen amma itti fayyadamaa jirru kanaan Afaan Oromon akka barra'uu heeran kan murtaa'ee bara 1991 turee. Qubeen Afaan Oromoo qubee barreeffama Afaan Oromoo akka ta'uu ega murtaa'ee boodas namooni adda addaa qubeen kun Afaan Oromootif mijataa akka hintaane qorannoo adda addaa taasisaniruu. Kun garuu akka dhunfaattii qorannoo gaggeessunkan danda'amu akka jirutti ta'ee haalli barreeffama Afaan Oromoo seeran murtaa'e waan ta'eef kan fooyya'uu qabus qoratamee, ummannii yaada irratti kenneetii maalee namoota dhunfaatiinii miti. Qorattoonni dhunfaan yeroo garaa garaattii haala barreeffama (qubeessutiin) walqabatee qorannoo adda addaa gaggeesaniru. Qorannoon kunis kanmul'isuu haala qubeessun walqabatee garaagarummaan jiraachu isaati. Haala qubeessutiin walqabatee yaadni ka'uu danda'a, duriis ka'aa turee waan ta'eef. Kana furufis koreenwaaltina Afaan Oromo hojjachaati jira. Qubee Afaan Oromootiin walqabatee garuu qorannoon koreen waaltina Afaan Oromootiin gaggeeffame hin jiru.

P1: Regarding Orthography and orthographic reform, article five of the state government of Oromia states that official working language of the state is Oromo and it should be written in Roman script and this was decided at the beginning of the establishment of the state government of Oromia. Before it was officially stated in the constitution of the regional state of Oromia, there

was a discussion by various scholars as to which script is favorable to write Oromo. Before 1991, Oromo was written by various authors using various scripts like using Roman, Arabic and Ge'ez scripts. Oromo scholars like Onismos Nasib, Aster Gano, Mohamed Rashid, Hayle Fida, Bekri Saphelo and political organizations like Oromo liberation front (OLF) have played a great role for the basement of the current script and orthography of Oromo. Just like the previous Oromo scholars did their best before the decision of the current script, various scholars like those who are from abroad, in the country and from various sector areas like education sector and other sectors have discussed as to which script is favorable to write Oromo. Various researches were conducted and presented to the Oromia state government as to which script is appropriate for writing and enriching the language.

P3:- Various scripts like Roman, Arabic and Ge'ez were suggested to write Oromo. It was after a long discussion that the present Roman script, called 'Qubee' became the official and the legal script for writing Oromo in 1991. Even after the Roman script was officially accepted to write Oromo, various individuals have continued claiming that the Roman script is not favorable for writing Oromo. If reform is needed for practical reason at all, the official script which was decided legally has to be studied and reformed after conducting a research and presenting it to the speakers and to the government. Various researchers have conducted a research regarding the orthography of the language and the findings of the study shows there exists a variation of orthography in Oromo. In order to solve the problem, the committee of standardization of Oromo is working. But regarding script 'Qubee', we did not conduct any research.

23. Q: Qubeen afaan Oromo amma ittin barreffamaa jiru Afaan Oromoof mijataadha jettanii yaaddu?

R: Do you think that the current working script of Oromo is suitable to write Oromo?

24. H4: Qubeen Afaan Oromoo amma ittiin barreessaa jirru kun Afaan Oromootiif baay'ee mijataafi afaanichis yeroo gabaabaa keessattii sadarkaa guddarra akka ga'u kan taasise ta'usaati. Kun ammo kanteef afaanicha salphaatti barachuuf waan tolfii. Dheeresanitis ta'ee gabaabsani salphumatti barreessuf mijataadha. Ummanis mijataa miti jijjiira nuun hinjenne. Kun ammo kan agarsiisu ummanni akka fudhatedha. Yoo haala qubeetiin walqabatee rakkinni kan jiru taanaanis

qorannoon gaggeeffame, qorannodhaan mirkanaa'ee motummaaf dhiyaatee ega ummanni fudhatee booda kan fooyya'u qabu.

P4: What we know is that the current Roman based script of Oromo is appropriate for Oromo. The current script is also suitable to easily mark germination and vowel length. It also enabled the language to grow faster as it can be learned easily. The speakers did not claim that this script is not appropriate to write Oromo. Hence, this can mean that they have accepted it. If there are script and Orthographic problems, a research has to be conducted and presented to the government and to the speakers (speech community) as a whole to be accepted.

25. H5: Akkuma sila jedhamettii tartiibnis ta'ee haalli qubeessun Afaan Oromoo hanqinalle yoo qabaate, qoratamee, saayinsidhan deegaramee, motummaaf dhiyaate ummataa fudhatamaa yoo argatedha kan fooyya'uu maale namoonni tokko yookin ammo qaamni tokko duwwaan fooyyeessuu hindanda'u; heera faanaa walqabata waan ta'ef. Dhimmii afaani heera fi imaammata wajjiin waan walqabatuuf kan fooyya'us heerani. Biirtoon barnoota Oromiyaa garuu gocha kana wajjiin walfaallessu hojjate. Haala tartiba qubee qorataniin yoo rakkinni tartiiba qubeetis ta'ee haala qubeessun walqabatee jira ta'e qorannoo sana qaama dhimmi ilaallatuf dhiyeesun akkuma jirutti ta'e, qorannoon gaggeeffame qaama dhimmi ilaallatuuf otoo hindhiyaatiin, otoo motummaan hin beekiinfi otoo ummanni irrattii mari'atee fudhatama hin argatiin tartiiba qubee jijjiruun yakka guddaadha.

P5:-Like it has been mentioned earlier, if there are the script and orthographic problems of Oromo, it has to be reformed after it has been studied and accepted by both speakers of the language and state government of Oromia. Nobody or no individual has a right to make any reform without the consultation of the concerned body as it has been decided legally. Any language related reform and decision has connection with legal decision and hence, if there is a need of language related reform, it has to be legally reformed in consultation with the concerned body of the government. But, education bureau of Oromia regional government did contrary to this requirement. If there are alphabetic and orthographic related problems, it has to be supported with various researches and the result of the study has to be presented to the concerned body.

But, changing the order of the alphabet or making a reform to the already accepted alphabet order without consulting the speakers and the concerned body of the government is an illegal act.

26. H2:- Heerrii motummaa naanno Oromiyaa keewwata shanirratti tartiibnii qubee Afaan Oromoo 'A, B, C,...' jedhee lafa kaa'erawaan ta'eef kanaan faallaa ' L, A,G,I,M...' jedhani tartiiba namnii itti wareegame jijjiruun ammo heera diiguudha. Motummaan naanno Oromiyaatis tartiibnii qubee jijjiramu isaatillee hinbeeknee. Ummannilee kan waan hinbeekneef balaan ka'eeti turee. Biroon barnoota Oromiyaatis dogoggoora kana raawwate. Haangoon biiroo barnoota Oromiyaatiif kenname afaanicha barsiisu malee tartiiba qubee jijjiirus ta'ee foyyeessu miti. Afaan Oromootiin hojjadhaa jedhaman malee Afaan Oromoorrattii akka hojjatan labsiidhaniis ta'e heeraan hinkennamnef. Biroon kun tartiiba qubee kan jijjiree koorpasii guureti qubee kamtu baay'inaan firiikuweensii baay'ee qaba kan jedhu qorataniis tartiiba qubee lafa kaa'ani. Tartiba qubee jijjirameenis kitaabni barataa maxanfamee waggaa lamaaf ega tajaajilee booda ummataan waan fudhatama dhabeef jijjirameera. Nutis ta'ee motummaan naannoo Oromiyaa quba hinqabnu. Isaanis nuniis ta'ee motummaa wajjiin hin mari'anne.

P2:- The alphabet order of Oromo has been stated under article five of the constitution of the regional state as 'A, B, C, D, E, F...' not as ' L, A, G, I, M,...'. Changing this on which the speaker lost their life is illegal. Even the concerned higher body of the regional state government and the speakers did not be aware that the reform has been made and there were strong oppositions from the speakers' side. The right and responsibility of the education bureau of the regional state of Oromia is to teach the language; but is not given by the law to make any reform and the bureau has committed this mistake. Education bureau of Oromia first collected texts and forms a text corpus. Secondly, the frequency with which each character used to write a word has been measured and order of the character has been set. Based on the reform, education bureau of Oromia published textbooks. But, the textbooks were not used more than two years as the speakers did not accept the reform made. Even we and the higher regional government bodies were not aware about the reform.

27. H1:-Tartiibni qubee jijjirame yaada ummataa, daddarbuu dandeettii fi ulaagaalee kan kana fakkaatan yaada keessa hingalchine. Tartiiba qubee jijjiruf ulaagaan ilaalamuu qabuu ammoo

baay'inaan irra deddeebi'ee argamuu qubee qofa otoo hintaanee, haala barumsaaf toluu isaatiifi, kan kana fakkaatan ilaallamuu qabu turan. Siilaa haala qubeessutiinis ta'ee haala tartiiba qubeetiin walqabatee qoratee qaamaa motuummaa dhimmii ilaalatuuf kan dhiyeessuu qabuufi labsidhan kan kenname Waaltina Afaan Oromoof ture.

P1:-Moreover, the reform did not take the acceptance of the speakers, transfer of skill and other criterion into consideration. Taking the frequency of the character is not the only criterion for setting the order of the character. Considering its suitability for teaching and learning of the language and other issues must be given due care. The right and responsibility of working on the language and other related issues in general as well as orthographic related issues in particular has been given to standardization committee of the language. Hence, this body is the right body to make any language related reform in consultation with the speakers and the concerned highest body of the state government.

28. Q :- Rakkinni haala qubeefi qubeessuu Afaan Oromoon faana walqabatee yoo jiraate otoo ibsitani ?

R:- Please discuss if there are any orthographic representation related problems in Oromo?

29. H1:- Akkuma kanaan dura ibsetti rakkinni qubee fi qubeessun walqabatee jira taanan furmaata kan argatuu qorannaan gagafamee, motummaafi ummataaf dhiyaatee fudhatama yoo argatedha heeran kan furamuu. Qubee Afaan Oromotiin walqabatee, garaagarummaa haala barreffama qubee chimdiirratti mul'atuus jiraa. Fakkeenyaaf qubeen 'ny' jedhu haalli barreeffamasaa qubee cimdiin kaniin wal hinfakkaatan. Qubeen cimdiin kuun dhumarratti 'h' qabu. 'Ny' ammo dhumarratti 'y' ofkeessaa qaba. Qubee Afaan Oromo salphaatti baruudhafis otoo walfakkeenya jiraatee gaaridha. Kunis akka amma ati qorattu kana qoratame motummaaf dhiyaatee ummanni yoo fudhatee heeranis kan mirkanaa'uu qabu. Itti fayyadama dhamsaga jechoota ergisaa wajjin dhufanirraattis garaagarumaan jira. Fakkeenyaaf jecha 'television'fi kan kana fakkaatan berreessuf miidiyaaleenifi namoonni garaagaraa haala wal hinfakkaanneen barreessan jiru. Kun 'zy' kun 'zy' kun 'j' jedhee barreessa. Garaagaruuman kan dhufes garaagaruma haal sagalessudhaani. Dhamsagni kunis gara fulduraattii qoratame qubee Afaan

Oromoo keessatti galuu yoo danda'ee mallattoo garaagaraatin otoo hintaane mallatto tokkoon bakka bu'uu qaba. Akkamitti akka barraa'u qabus murtaa'u qaba. Furmaatas argachuu qaba. Baay'inni qubee dubifamaa Afaan Oromo qorataadhaa gara qorataatitti adda adda. Kun baay'inni isaani 22, kun 25, kun 27, kun 28 kun mana soddomaas kan galchan jiruu. Nutii garuu digdamii saddeet jenna.

P1. Like it has been stated previously, if there are any orthographic problems, it has to be solved after conducting and presenting a research to the speakers and to the concerned body of the state government. It has to also be accepted. There is a difference in orthographic representation of digraphs. For instance, the digraph ch, dh, sh, ph contain 'h' at the end of each digraph. But the digraphny contains 'y' at the end. Learning the alphabet of the language will be easy if such uniformity problem is corrected. This also has to be studied like you are studying it now. The study has to be presented to the government and to speakers and be accepted. There is also a phoneme called /ʒ/ which entered through loanwords as in 'television'. Various individuals and mass media have represented this phoneme differently. This phoneme is represented using different graphemes as Zy, Zy, J. This is due to the fact that we did not still either represented the phoneme with grapheme or made any decision officially. So this has to be studied and either has to be represented with agrapheme or has to be clearly stated as how such kind of loan words having the phoneme /ʒ/ should be written in the future. The number of consonant phonemes of Oromo is different from author to author. Some of them make it, 22, and other make it 25, and still there are authors who claims the number to be 27 and 28. But we claim the number to be 28.

30. H4:- Karaa biraatiin ammoo hudhaan qubee dubifamaafi dhamsaga afaanichaa ta'ee fudhatamee otoo jiru tartiba qubee keessatti otoo hinramadamin hafuunsaa qoratamuu qaba; qorachuus qabna. Hudhaan qubee dubifamaa ta'ee mallatto hudhaadhan (') bakka bu'ee ittiin barreeffamaa otoo jiruu, iddoo baay'eettis otoo argamuu, tartiiba qubee keessatti hin ramadamnee. Kun kan ta'eef ammoo hudhaadhaaf mallatoon qubee waan hin boocamneef. Hudhaan dhamsaga Afaan Oromoo ta'udhan alaatti haala qubeessu Afaan Oromoo keessattis baay'inaan kan argamu waan ta'eef qoratamee akka furmaata argatuuf nutis hojjachu qabna. Hudhaan yeroo tokko tokko 'w, y fi h' dhan bakka bu'ee barreeffamaa jira. Hojiin afaan waaltessufi qorannaan afaani waggottan baay'ee fudhata. Kanaafu kunis qoratamee heeran

furmaata argachuu qaba. Bara 1991 akka lakkoofsa warra Awrooppaatti sagaleen hudhaan akka jecha tokko giduutti qofa akka barraa'uu murtaa'ee ture. Bara tokko ammoo sagalumtii kunjecha giddurratiis ta'ee yeroo jechii qubee dubbachiftuu dheeraadhaan yoo jalqabee akka sagalee hudhaan mul'atuu (durarratti barreeffamu) murta'eet ture. Murteen kunii garuu yeroo baay'ee otoo hinturiin akka murtee duraatitti deebi'uu afaaniin darbeera. Murtoonni akkanaa workaan hin barraa'ee ummataaf hinga'iin ammoo waltina afaanichaarrattii dhibaa guddaa umma.

P4: On the other hand, though glottal stop /ʔ/ has a phonemic status and included in the phonemic inventory of the language, the phoneme is not yet represented with a grapheme and it is not included in the alphabet of the language. This has to be studied and we are the concerned body to study it. The glottal stop is one of the most frequently occurring phoneme in the orthography of the language which is being indicated by an apostrophe ('). Since we have included in the phoneme in the phonemic inventory of the language, we have to study and find out how it could be represented using by grapheme. Sometimes this phoneme is also being seen replaced and written by 'w, y fi h'. Standardizing the language takes several years. So, this has to be studied and the solution has to be given legally. In 1991, the glottal stop was decision to be applies only word medially; but some years later there was also a convention that states in addition to word medially, it has to be used word initially when a word begins with a long vowel. But, soon, the decision was reversed into its original convention orally without stating it in any document of the authority. I think, this will definitely affect the standardization process of the language.

31. H5: Qubee Afaan Oromootiin walqabatee akka rakkotii otoo hintaane akka yaadattii tartiibnii qubee amma ittin barachaafi barsiisaa jirru quboota cimdii dhumarratti barreeffaman akka barumsaaf toluttii gara 'A, B, C, Ch, D, E, F, G, H, I, J, K, L, M, N, Ny, O, P, Ph, Q, R, S, Sh, T, U, V, W, X, Y, Z jedhuttii otoo jijjiramee. Kun ammo nama Afaan Oromoo akka haaraattii baratuuf tartiibasaanifi booca saanii salphaatti barachuuf tola.

P5: Regarding the alignment of Oromo alphabet, it will be easy to learn the language if it is rearranged as 'A, B, C, Ch, D, E, F, G, H, I, J, K, L, M, N, Ny, O, P, Ph, Q, R, S, Sh, T, U, V, W,

X, Y, Z'. This will make learning easy especially for beginners. This will help to learn the digraphs easily.

32. Q:- Adeemsa waaltessuu Afaan Oromoo keessatti loga (logoota) kam bu'ureffattan?

R: Which variety (s) is (are) being used as base dialect in codification processes of Oromo?

33. H1:- Afaan waalteessuu kessaatti logoota jiran kessaa looga kamtuu akka bu'uratti fudhatamuu qaba kan jedhuuf debiilee adda addaatu kennama. Innii tokkoffaan loguuma jiraan keessaa midiyaalee, baranootafi tajaajila adda addaarratti kan ittii hojjatamaa jirrufi iddo bal'attii kan fayyada jiru filatamuu danda'a. Kun looga bu'uraa tokkee 'mono centric' jedhama. Lammaffaa loogoota hunda akka bu'ura fudhachudhan 'polycentric' afaan waaltessuudha. Afaan Oromo waaltessuf koreenifi gumiin waaltina Afaan Oromoo filanno lammaffa kana fudhatee waggaa digdamii shaniif ittin hojjachaa ture. Looga tokko qofa fudhatanii waaltessuuf yaalun balaadha. Qaamni waaltessu suunis looga tokkotti looguun tari jiraachuu danda'a. Kun ammo balaa fida; godina afaanifis rakkina fida; kessumaayuu waaltessu kessattis rakkiina fida. Rakko kana furuf loogota hunda bu'ureffannee, isa 'composite dialect' jedhamun waaltessaa jirra. Fakkeenyaaf nutii akka koreefi gumii waaltina Afaan Oromootitti jechootas ta'ee wanta waalta'u qabu yeroo waaltessinu hiika sana loogoota adda addarrattii maal fakkatu jenne ega barreesinee boodaa kamtuu filatamu qaba kan jedhu murteesina. Kan afan ammo akka 'variant' tittii lafa goona. Kunis yeroo baay'ee boodavaryaanittin tokko caallatiifi baal'inaan itti fayyadamni isaa yoo jiraatee jijjirree isaa waaltessina.

P1: From the conceptual point of view, there are various options regarding the choice of the base dialect from the existing variety of the Oromo. The first option is elevating a sing variety which is currently being used in the mass media, in the education and in the various spheres. This is the variety being used widely. Such kind of choice of base dialect is said to be a monocentric. In order to standardize Oromo, a monocentric approach of base dialect choice is very dangerous. It will also be a problem for the growth and enrichment of the language. We consider this approach to be difficult to standardize the language. The second approach is taking all varieties of the language as base dialect and such an approach said to be a polycentric or composite approach. Standardization committee has chosen the second option of base dialect to standardize the

language and we have been working for the last 25 years. When we codify lexicon, we first see and collect how that word is written and represented in various varieties of the language and then we will discuss about which variety has to be taken as a standard. We will put others as variant for that lexicon. If one of the variants of which previously not taken as standard become dominant as the time passes, then we will change the previous one with the new dominant variant and it will be taken as standard. We will put the previous as variant for that specific lexicon.

34. H2:- Sababni koreen waaltina Afaan Oromoo dhaabbateef garaagaruummaa barreeffama keessatti mul'atuu waaltessuufi. Loogoota jiran keessaa tokko mosiisuuf miti. Loogoota jiran sakkatteesinee looga jidugala fudhannee waaltessaa jirra. Kanaafuu jechootas ta'ee latiwwaanfi kan kana fakkaatan yoo waaltessinu loogaa garagaraarraa fudhanneetuu.

P2. The main reason for establishment of standardization committee is to reduce variation and to standardize the language, particularly the written aspect and not to make one of the variants superior. Hence, while standardizing lexicon, morphemes and other issues, we will, first, collect from various variants, then we will decide which variant to be codified as a norm and the remaining will be there as variant of the selected norm.

35. H6:- Afaan Oromoo akka afaan biraattitii loogoota jiran keessaa looguma tokko filannee waaltessuf baay'ee rakkisaadha. Loogoota hunda fudhannee looga gidugalaa bu'ureeffanee haga ammattii waaltessaa jira. Kunis isa 'composite base approach or polycentric approach' jedhamu sani jechuudha.

P6:- It is very difficult to standardize the language through choice of only one of the variants and elevating it as a supradialectal norm like other languages. We are standardizing the language taking the composite approach of language standardization.

36. Q:- Loogaa gara loogatti dhaaba dhabuun (addaddummaan jechootaa) yoo jiraatee otto ibsitani? Yoo garaagarummaan jira ta'e, afaanicha waalteessuu keessatti rakkoo jiruu otto ibsitani? Garaagarummaa kana furuuf maaltu hojjatamaa jira?

R:- Would you please discuss if there are lexical variations in the language? What its challenges in standardizing the language? What is being done to solve the variation?

37. H3:- Koreen waaltina Afaan Oromoo bara 1985rraa jalqabee hangoon afaanicha waalteessu waan kenameef haga ammaatittii waggaa 25f afaanicha waaltessa jira. Haala sagaleessurraa jalaqabee garaagarummaan jiru xiqqeechuf hojjachaa jirra. Biroon aadaafi tuurizimiin akka hojii kana hojjatu labsii 198/2008dhaan kennameefira. Kun ammo danbiidhan koree waaltina Afaan Oromootif qaqqabeera. Afaan kana jidugaleessummaan akka waaltissu heeran haangoon keenname jira. Hojiirra olmaasaatis nitoo'ata. Bakkee rakkinni jiruttis sirreffama ni keenna. Afaan Oromoo keessattii jechooni loogaa gara loogatti garagarummaa qaban baay'ee bal'aadha. Haala qabatama haala qubeessudhaanis ta'ee haala ittii fayyadama jechootaatiin garaagarummaa heedduu jira. Fakkeenyaaf, 'ogbarruu' barreessuuf kuun 'og-barruu' kuun 'hog-barruu' kuun 'hogbarruu' jedhee barressa. Isa Ingiliffaan 'egg' jedhuuf'*haanqaaquu, buphaa, killee*', isa 'false jedhuuf 'saba, dhara kijiba' jedhamee haala sadiin kan barreffamu. Kun ammo afaanich gabbisuufis ta'ee afaanicha baruuf dhibaa guddaa qaba. Kan akka fakeenyatitti kaafne maale garaagarummaan jiruu kana qofa mitii. Kan akka kanaa barreeffama garagaarrattii ni hubatamu. Kana furufis hojjachaa jirra.

P3:- It was in 1991 that the standardization committee of Oromo was given legal responsibility to standardize the language. Therefore, we have been working to standardize the language for the last 25 years. We are working to minimize the variation. Culture and tourism bureau of the regional state of Oromia has been given tasks of working on the language on proclamation number 198/2008. The bureau has given a legal responsibility to us through regulation and we have the mandate to work on the language and minimize variation. There are wide lexical variations among the variants of the language. The variation extends from Orthography to the grammar. For instance, the word 'literature' is written as '*Og-barruu, Ogbarruu, hog-barruu, hogbarruu* oguma'. There is also difference for the concept 'egg'; '*killee, buphaa, hanqaaqu*'. The word 'false' has three variants; '*soba, kijiba, dhara*'. The variation is very numerous and what I have given is just to show instances of variation. The variation can affect the growth and enrichment of the language; particularly standardizing its written aspect as we have observed

such variation in the various texts. It has also an effect on learning the language. We are working to minimize the variation.

38. H4:- Afaan Oromoon keessumaayyu garaagaruumaa jechoota fi latiiwaani baal'inaan qaban keessaa afaan isa tokko. Karaa biraatiin ammo namoonni adda addaa yuunivarsiitii ta'ee warrii kitaaba barataa sadarkaa tokkoffaafi lammaffaa qopheessan keessa jiran illee jecha haaraatiif nutii waan moggaasne dhiisani haalaa garaagaraatiin boocaa jiru. Fakkeenyaaf isa 'adverb' jedhuf bakka adda addaatti ' ibsa xumuraa, gochibsa, xumibsi' jedhanii yoo boocan, koree ammo 'ibsa xumura' jedhee moggaase. Nuti jechoota moggaafnee waaltessine dhiisanii ofiif boocachaa jiru. Kun ammoo isaan hojii afaanicha waaltessuun faallaatti hojjachaa akka jiran hubanna. Kana furuufis koreen waaltina Afaan Oromoo jechoota moggaassee waaltessite gidugala tokkorraa akka tamsa'uu waggootan 25f moggaasnerra; kun ammas itti fufaa jira.

P4:- Oromo is one of the languages having wide variations, particularly in relation to its lexical and morphological variation. On the other hand, to the contrary of using words that we have already codified, various individuals; particularly those who are from various universities are coining new words differently which may play against standardization of the language. For instance, for the word 'adverb' various authors has codified it as ' ibsa xumuraa, gochibsa, *xumibsi*' and the committee codified it as '*ibsa xumura*'. In order to solve such variations we have been codifying lexicons for the last 25 years and we will continue to codify.

39. Q:- Yaada haaraatif Afaanuma Oromoo keessaa akkamittin jechoota moggaafu? Yaada kanaan dura moggaafan tokkof jechi caalatti ibsuu yoo hargattan jijjirtani beektu?

R: How do you coin new words for new concepts from Oromo? Do you have ways to change the codified words whenever you get better words for the concept?

40. H5:- Yaada haaraatif Afaanuma Oromoo keessaa jechoota haala adda addaatiin moggaafna. Fakkeenyaaf jecha lamarraa jecha tokko boocuu (xumura jecha duraa irraa qubee dubbachiftuu yookan dubbifamaa yookan lamaanu hanbisudhan) kan akk '*fakkoo + Ardii*→*fakkardii*', '*faaya + yeroo*→*faayeroo*' garaa biraatiin jechoo lamarra dubbifamaa fi dubbachiftuu muruudhanis ni

boocna. Fakkeenyaaf 'afaan+ dhabaa→afdhaba, 'afaan + makaa→afmaka' fi kan kana fakkaatanin jechoota boocna.

P5:- We codify new lexicon for new concepts and ideas from the language itself using various techniques. For instance, we codify new lexicon from two different words i.e taking only some part of the first word and attaching it into the second word like a prefix.

For instance, 'fakkoo' to mean 'picture'
 'ardii' to mean 'content'
 fakkoo + ardii→ fakkardii 'planet'

faaya 'close' + yeroo 'time' →faayeroo

41. H3:-Yaada haaraatiif Afaanuma Oromoorraa tooftaa adda addaattiin boocna. Jecha hiika tokko qabuu hiika dabalata akka qabu gochuun 'semantic extension'. Fakkeenyaaf:

gocha	'action'	'action verb'
fuula	'face'	page
kofoo	'trousers'	angle
ijaa	'eye'	important
afaan	'mouth'	mouth

Yaada kana dura moggaafnef tokko jechi caalaatti ibsu yoo arganne jijjirrees beekna; nijjjiras

P3: We codify lexicon for new concept through various techniques. Semantic extension is one of the techniques we use. The following illustration shows semantic extension technique of lexical elaboration in Oromo.

Oromo word	meaning1 (Literal)	meaning2 (extended)
gocha	action	'action verb'
fuula	face	'page
kofoo	trousers	'angle'
ijaa	'eye'	'important'
afaan	mouth	'language'

42. Q:- Jechoota afaan alaarraa gara Afaan Oromootittii toftaawaan itti madaqsitan otoo ibsitani?

R: Would you please discuss the lexical elaboration techniques you frequently use to designate foreign concepts in to Oromo?

43. H1: Jechoota afaan alaarraa gara Afaan Oromootittii toftaawaan itti madaksinu kessaa anni tokkooffaan jechii madaqfamu caasaa afaanichaaf akka toluu gochuun eergifachudhan ‘modified borrowing’ Fakkenyaaf: Jechoota afaan biroorraa otoo hin jijjiriin akkuma jiruttii fudhachuudhan, 'boorsaa' afaan Xaaliyaaniiraa kan fadhatamee ‘borsa’ ‘bag’, 'kartaa' ‘carta’ ‘map’. Inni lammaffaa seera Afaan Oromootti madaqsuu ‘Addaptation’

- piroojaktii
- baankii
- saayinsii
- ispoortii
- kompitara

kana akka fakkeenyatitti kaafne malee tooftaan afaan alaarraa ittiin madaqsinu kana qofa miti.

P1. There are various ways of lexical elaboration techniques we use to designate foreign concepts in to Oromo. Some of such instances are the following. For instance, take the word with modification (with vowel length).

Borrowed word	Oromo	gloss
borsa	boorsaa	bag
carta	kaartaa	map

The other lexical elaboration technique is borrowing through adaptation. The following are examples of borrowing through adaptation.

piroojaktii	'project'
baankii	'bank'
saayinsii	'science'
ispoortii	'sport'
kompitara	'computer'

These are not the only lexical elaboration technique we use.

44. Q:-Jechoota isiin moggaaftaan fudhatama argachuu isaani akkamittii mirkanissitu?

R: Explain your ways of checking whether the codified words or terms have got an acceptance or not?

45. H6:- Garaagarummaa jirurratti haayyoonni qorrannoo gaggeesiniiru. Kora adda addaallee qopheesinee qorrannoon adda addaa kora kanarratti nidhiyaatuu. Nutis argannoo kana nifudhanna. Hanqinaalee jiran, rakkoolee jiran, kan haga ammaatti koreen waaltessee wiirturrattii baafna. Dhaabbilee garagaraa wajjinis marii ni gaggeessina. Dhaabbilee adda addaaf ni raabsina. Dhaabbilee adda addaafi yuunivarsitii adda addaa demnee jechoota nutii moggaafnee ammam fudhatama akka argatan yeroo hundas ta'uu batuu horidoffii nigoona. Garaagarummaan dhaabbileefi nu giddu jiruu qorrannoo gaggeessine sirreeffama kan barbaadu sirreesine beekna nisirressinas. Kana yoo jennu garuu hanqinnis akka jiruu baruun barbaachisaa dha. Kan waaltisinerra kan hinwaaltofnetuu caala. Haala qabatama kanaa olitti caqafaman ilaalun hojiin waaltinii afaanichaa gara fuldurattii maal akka ta'u qabuus nikaroorfanna.

P6:- There are individual researchers who were conducting a research regarding the existing variation. The findings of all of these researches will be presented during the research conference we organize. We consider the findings of all those research and evaluate whether codified norms are being accepted or not. After codifying and distributing to various bodies, we will check their practicability in the various institutions and in the various spheres. If there is a need that the codified words have to be changed, we will reconsider it. This does not mean that there is no problem at all as there are more issues to be standardized than the one we already did. Taking all those issues into consideration, we will plan what to standardize in the future.

46. H4:- Yaada haaraatiis jechoota kan mooggaasnu loogoota siirriittii sakkataanee yaada sanaaf jecha yoo dhabne. Jechoota nutii moggaasnee raadiyoo fi televijiinidhaanis yeroo ummannii tamsaasa kallattiin yaada akka itti keennu goonus jira. Kan jechoota moggaafne maxxansinee yuunivarsititti raabsinee saanaaf isaan ofiisaanitiif akkamitti akka moggaasanii ittin fayyadamaa turanis ni gaafanna. Kan waaltessine otuu jiru kan isaani maalif akka fayyadamanis nigaafana. Kan waaltessinerrattis yaada akka kennan yeroo goonees jira. Itti fufeenyarrattii garuu hanqinni jira.

P4: We codify lexicon for new concepts after we make sure that there is no word from the existing variants that can express the new concept. We sometime sent the codified lexicon to the media to be evaluated by the speakers through live transmission. We also sometimes compare what we codify with what the higher institutions (universities) codify for the same concept. We have also asked why they prefer to codify their own than using what we have already codified for the new concept. We also sometimes ask them to comment on what we codify.

47. Q:- Jechootaa afaan biraa irraa ergiffachuun dhirqama yoo ta'ee, jechoota afaan kam irra caalaa filattu? Maalif?

R:- If lexical borrowing is the final option, which language(s) do you prefer and why?

48. H1:- Dura jechoota kan boocnu afaanuma keenyarraayi. Kan ergiffannu ergiffachuun dirqama yoo ta'ee qoofa. Jechoota kan ergifannu afaanuma jechii madderra ta'ee haala seera afaanichaatti madaqsudhaani. Dura kan ergifannus afaan biroo biyya keenyarraayi. Kunis namooni afaan biyya keenyaa beekan salphumatti jechoota sana akka barataniif.

P1:- We prefer to codify from the language itself. We first prefer to borrow from other languages in the country if there is no option from the language itself. This is due to the justification that those speakers of other Ethiopian languages can easily learn the borrowed words as the words are taken from their language. There are also cases where we borrow words from the source language with some modifications.

49. H5:- Kan dura sakattaanu afaanum nu dhiyoo jiruu. Afaan biyya alaa keessaa otoo hinfudhatiin afaanuma ollaa keenya keessa jiru ilaalla.

P5:-We, first search from the neighbouring languages before borrowing from foreign languages.

50. Q:- Jechoota moggaafan akkamitti uummata biraan geessuu?

R: How do you disseminate the newly coined lexicons to the users?

51. H1:- Kora qopheesinerratti raabsina. Kan waaltesinee ummataaf ni raabsina. Manneen kitaaboota sadarkaa tokko haga yuunivarsititiitti jiraniif, barsiisoota sadarkaa tokkoffaa haga yuunivarsiitii jiranif, midiyaalee filataman adda addaatiif niraabsiina. waggaa sadiin dura jalqabee garuu hanqinni raabsuu nijira. Komiinis dhayaachaa jira. Dura garuu godina hundarra deemne raabsaa turee.

P1:- We have distributed the codified words to various bodies via various options; the distributions of the codified words during the research conference can be taken as an example. We also reach the speakers through distributing to the public and schools (which extends from primary to universities) libraries and to some selected media. But there is a distribution problem starting from three years back to now. Due to this problem, there are various questions from the speaker side. But before three years back, we used to distribute to each zone.

52. H3:-Waan waaltessinee maxansinee midiyaaleetiif, biiroo barnootaatiif, yuunivarsitif, mana barumsaa sadarkaa tokkoffaafi lammaffaatiif raabsina.

P3:- We distribute the codified words to the mass media, to the education bureau of Oromia and to various levels of education institutions.

53. Q:- Afaan Oromo waalteesuun walqabatee haga ammatti qaamoota dhimmii ilaallatuf (fakkeenyaaf, ogeeyyi barnootaa fi ogeeyyii sub qunnamtiif leenjii kennitanii yookaan waliin marii gootanii beektuu?

R:- Have you conducted any training or discussion about standardization of Oromo for stakeholders like education and media expertise?

54. H1:- Dhaaboota kana ol caaqaftee wajjiin kopha kophaa isaanii wajjiin mari'achuu yoo baanneellee kora adda addaa qopheessinurratti nihirmaachisna. Yeroo sana waliin nimari'anna, haanqina jirus nifudhana. Naannoolee biyyatti kana keessa jiran keessaa naannoo Oromiyaa qofa kan koree waaltina afaanii hundessee waggaa digdamii shan ol hojjachaa kan jiru. Miseensonni keenyas, dhaabilee barnoota, subqunamitii adda addaa (kan akka OBN, ETV, FBC), dhimma

kominikeeshinii motummaa faadha. Ciminaan hirmaachurratti rakkinnii nijira. Keesumaayu biroo bornoota Oromiyaa wajjin waltii hiqnee hojjachurrattifi kitaabni barataafi galmeen jechoota yeroo qophaa'u walqabatee waliin hojjachurratti hanqinni jira. Biroon kitaaba barataas ta'ee galmee jechoota yeroo qopheessu nu wajjiin hinmari'atu.

P1: We invite those bodies mentioned above to participate on the various research and discussion related conferences we prepare though we didnot make a training and a discussion with each of them individually. We always consider the gaps stated in the conference. It is only Oromia regional government that established language standardization committee and council to standardize the language and it has been working for the last 25 years. Education bureau of Oromia, electronic broadcasting media (particularly OBN, ETV, FBC) and the communication affairs of the government are working with us. But there is gap regarding working together continually. Particularly, there are gaps of working in collaboration with the education bureau of Oromia. They do not even consult us when they prepare dictionaries and textbooks.

55. H3: Afaan Oromoo waaltessuuf koreen waaltina Afaan Oromoo guyyaa guyyaadhan hojii afaanicha waaltessuurratti hojjachaa jira. Torban lama lamattiis walga'eedhan waan waaltessine sanarratti marii ni godha. Qaamni itti anuu kan barana hundaa'ee gumii waaltina Afaan Oromoo jedhama. Gumiin kun ji'a sadi sadiittii walga'ii geggeessu. Miseensii gumuu kanaa midiyaalee (kan akka OBN, ETV, FBC, Kallacha Oromiyaa), biroo barnoota Oromiyaa, haayyulee yuunivarsitittiifi sadarkaa lamaffaattii barsiisan, haayyulee afaanii looga hundarraa filataman, gaazeexesitoota, biroo dhimma kominikeeshinii motummaa fa'i. Kanaa ala ammo qorattootaadda addaa hawwachudhan simposiyamii adda addaas ni gaggeessina. Leenjiiidhaan walqabatee garuu wanti nutii geggeessine hinjiru.

P3: The committee is performing its task every day to standardize the language. The committee also meets after each two weeks to discuss about what we have worked. The standardization council of Oromo, the newly established next higher body of standardization of the language, meets after each three months to discuss and comment on standardization of Oromo. The media mentioned above, Kallacha Oromia (a print media), education institution, some selected university lecturers, media expertise, education and government communication affair, and those

language expertises from each of the varieties of the language are the member of the standardization council of Oromo.

56. Q:- Haala barreeffama Afaan Oromoon walqabtee rakkooleen kitaaba barataa sadarkaa tokkoffaaakkasummas midiyaalee biroodkaastiwaan Afaan Oromoo keessatti waan argitan jiraa? Yoo jiraatee, garaagarumma jiruu furuuf maaltu hojjatamaa jira?

R: Would you explain if there are orthographic related problems being observed in the broadcast media and in the primary and secondary school text books? How do you solve?

57. H6:- Garaagarummaan haala qubeessuun walqabatee kitaaba barataa sadarkaa tokkoffaafi lammaffaas ta'ee midiyaalee biroodkaastii keessattii rakkooleen baay'inaan jiruu. Kanas sirreessufi waaltessu qabna; waalteessuun ammo angoo keenya. Jecha nuti waaltessine dhiisanii haala garagaraattiin kan barressaan jiru. Jechoota tishoo, dhamjechootafi akka waliigalaatitti haala qubessurrattii rakkinni jira. Fakkenyaf, nutii 'fi, dha, ti' jecha isaan dura jiru wajjin akka jecha tokkotittii waan dubifamanif jecharratti maxxananii haabarreffaman jenna. Kana 'manual' dhaan barressine ummataaf hinraabsine. Kun hanqina. Kitaaboota fi midiyaalee kanaa olitti caqafaman keessatti garuu qophaas barreessaa akka jiran argaa jirra. Kana furuf ammo dhaabbilee kanaa olitti eeraman nuwaliin hojjachaa jiru. Kana caalaas cimsinee hojjachuu qabna. Kan waaltessines raabsaa jirra.

P6: There are a number of orthographic variations in the textbooks of primary schools and in the broadcast media. They write differently without looking at what has been codified. Variation includes orthographic variation in writing compounds, morphemes, affixes, conjunctions and so forth. For instance <fi>, <dha>, <ti>are not seen to be written attached to words as they are pronounced as part of the word before them. But, we codify them to be written attached to the words; though we did not prepare a manual and distribute to the users. We have to correct (minimize) such kind of variation as we have been given legally to standardize the language. We have to work hard to minimize the variation.

58: H2:- Rakkolee kun kitaaba barataarratti, barjaaleerrattii, xalayaalee biroowwaniin barreeffaman keessa, kitaaboota adda addaafi barreeffama adda addaa keessatti nimul'atu.

Fakkeenyaaf akkuma sila jedhamettii 'dha, irra, ni, fi, kkf yoo barressan kuun jechatti maxxanseeti barreessa kuun kophaatti barreessa. Garagarummaa baay'eetu jira. Nuti akka qaama afaan kana waaltessutittii qophasaanii kan hiika hinqabnee ittihiiqani haabarra'ani jenna malee akka afaan biroo kophaasaani haabarreffaman hinjennee. Bakka adda addaatti namooni adda addaa kan nuti waalteessine dhiisani kanisaani ummatani kan barressanis jiru. Kuun ammo lakkii kan waalta'een barreesina jedhani haala nuti waaltessineenis kan barressanis jiru. Gama keenyatinis rakkinni jira. Haala qubeessurratti 'guideline or manual ' qopheesinee iddo hundaan hingenye, kanas hin too'anne. Rakkoo kana furuufis qoranno barbaada.

P2: Various orthographic variations are being observed in the text books, advertise banners, letters written by various bereau, news and other texts. There are different orthographic variations like the one given previously. We codify that those which cannot stand alone should be written attached to the word and should not be written alone like morphemes of some other languages. Various individuals are writing differently leaving the codified one. To the contrary, there are also individuals who are committed to write using the codified one. We have also problem as we did not prepare and distribute guideline or manual. We did not follow up and control to the extent needed.

59. H5:- Akka jechoota waaltessa jirru kanatti haala qubeessu waaltessurratti hinhojjanee. Jechooni tishoo, walqabsiistuun, latileen danomiina agarsiisanfi kkf akkamitti barreffamu kan jedhurratti hanqina guddaatu jira. Gara fulduraa rakkoo kana furuuf qorannoo gaggeessu qabna; maanuwaaliis qopheesine raabsuus qabna. Itti fayyadamasaatiss hordofuu qabna.

P5:- We also did not work much on orthography like we worked on lexical codification. We did not set norms as to how compound words, conjunctions, morphemes, plural markers and others should be written. Conducting a research is needed here. We have to prepare a manual or guideline to solve such problems and then, we have to also follow up its implementation.

60. Q:- Afaan Oromoo akka gaaritti waaltessuf hojjiin kana caalaa hojjatamuu qaba jittan otoo ertanii?

R:- What do you suggest for further standardization of Oromo?

61. H1:- Dhugaadha. Afaan Oromoo kana caalaa waaltessuf kana caalaa hojjatamuu qaba. Afaan Oromoo nan gudisa, jettee waanuma haawwiteef Afaan Oromoon niguddata jechuu miti. Caasaasaa, haangoosaa, ogeessasaa guddisuu qabna. Akkaata amma hojjannuun gaafiin ummata Oromoo deebi'uu hin danda'u. Kana caalattii hojjachuuf dursa ammo ogeeyyii dhimmi ilaalu, muxannoo, dandeetti, saayinsiisaa sirritti beekan gara kanatti fiduun barbaachisaadha. Motummaanis sadarkaa caasaa kanaa yoo xiqqaate sadarkaa yuunivarsitii caalutti akka guddatuufi ogeeyyiis akka haawwatu gochuuf akka qorannee dhiyeessinu gaafateera. Nutis qorannee motummaaf dhiyeesineerra. Koreedhan afaan waaltessun hojjii koree ta'eetu hafa. Caasaasaa gara sadarkaa olaanaatitti guddisuun ni barbaachisa.

P1: It is true that in order to standardize Oromo, more has to be done. Wish alone does not make Oromo to grow; more has to be done. We have to change its structure to the next higher stage (at academy level) and there should be a professional and competent expertise. The Oromia regional government has given us to study its structure and to make it to the status; at least equal to or greater than the university level. This is to attract qualified language expertise that it requires. Otherwise, standardizing the language at the level of the committee does not bring the expected change. Hence, changing its structure and responsibility to the next higher position is mandatory.

62. H3:-Afaan Oromo kana caala waaltessuf wantoonni adda addaa fooyya'u qabu. Tokkoffaa akkuma jedhame caasaan amma ittiin hojjachaa jirruu kun gara akkaadaamittii ol guddachuu qaba. Waa'ee waaltina afaanichaarrattii qorannoo hedduu gaggeeffamu qaba. Haayyuleen amma jiranis gahaa miti. Hojiin waaltina afaanichaa gara fulduraa hojjatamuu qabuu humna amma qabnu wajjiin walhingituu. Ogeessa beekkamaa ofitti simachuu qabna.

P3: Various things have to be changed in order to standardize the language better. First, just like it has been mentioned earlier, the current working status has to be changed to the academy level so as to conduct various standardization related studies. The number of expertises has to be increased as there are a lot of standardization related issues that has to be normalized in the future and this could not be done by the existing expertise.

63. Q: Baay'ee galatoomaa!

R:- Thank you so much!

64. H:- Atille ulfaadhu!

P:- Thank you too!

II) Transcription of focus group discussion two (FGD 2)

1. Q. Kitaabnii Caasluga Afaan Oromo waalta'ee kan qophaa'efi akka wabiititti tajaajilu kan koreedhan yookan biroo aadaa fi turizimiitiin qophaa'e jiraa?

R:- Is there any standardized grammar book of Oromo developed by the committee or by Oromo culture and tourism bureau that is being used as codified norms for the users?

2. H7:- Kitaabni caasluga Afaan Oromo akka wabiitittii tajaajilaa jiruu biroo adaafi turizimiitiin kan qophaa'edha. Kitaabnikun barattoota sadarkaa lammaffaafi yuunivarsiitiif akka wabiitittii tajaajilaa jira; Kitaabni kunis yeroo lama fooyya'e.

P7:- There is a grammar book of Oromo developed by Oromo culture and tourism bureau that is used as a reference in the secondary and in the university and it has been revised two times.

3. Q:-kitaabni kun waalta'adha jettanii yaaddu?

R: Do you think that this book is a standardized book?

4. H8:- Akkas jechuuf rakkiisaadha. Ammas kitaabni kun rakkina waaltinaa qaba. Ammas irra deebi'amee fooyya'ee waalta'uu qaba. Kana boodas hanqinasaa laale kan fooyyeessuu qabus koree waaltina Afaan Oromooti.

P8:- It is impossible to say so. It has been edited and published for the second time and even this second edition has various codification problems and it has to be standardized again by standardization committee.

5. Q: Loogni Afaan Oromoo koreedhan yookan ammoo qaama mottummaa dhimmii ilaallatuun qoratamee qoqqoodamee beekaa?

R: Explain if you or any concerned government bodies have ever classified the dialects of Oromo?

6. H9: Loogni Afaan Oromoo qorattoota adda addaattiin bakka adda addaatitii qoqqoodamera. Akka koree waaltina Afaan Oromootittii garuu qorannee qoqqoodnee hinbeeknu. Gara fulduraa hojjachuf karoora qabna. Looga Afaan Oromoo qorachuuf waggaa afuriifi shan caala fudhata. Baasiini inni gaafatus salphaa mitii.

P9: Various researchers classified the dialects of Oromo differently. But dialects of Oromo have never been studied and classified by the standardization committee of Oromo as it requires lots of budget and time. We have a plan to study the dialects of Oromo in the future.

7. Q: Kitaaba caasluga isiiniin yookaan biiroo aadaa fi turizimiittiin barreeffame keessatti looga kam bu'ureffattan?

R: Which dialect(s) is (are) taken as base dialect in the grammar book you developed or the grammar book developed by Oromia culture and tourism bureau?

8. H10: Kitaabni caaslugaa bara 1998 A.L.E kan maxxanfame kan bu'ureeffate looga hunda. Looga tokko qofa moosifnee kitaaba caaslugaa kana hinbarreesine. Isa 'composite' approach' jedhamunii caaslugas ta'ee jechoota kan waaltessinu.

P10: The grammar book published in 2006 is compiled through composite approach. The publications have not been compiled through taking one dialect as the super dialect norm. It is through the composite approach that we codify the language.

9. Q: Kitaaboota caasluga Afaan Oromoo keessaa kitaaba kamtuu akka kitaaba waalta'etiitti fayyadaa jira? Kitaaboonni barataa sadarkaa tokkoffaa Afaan Oromootin barreeffamanifi oduuwwan biroodkaastiidhaan tamsa'an afaan waalta'een kan qopha'an isinitti fakkaataa?

R: Which grammar book of Oromo is used as a standard grammar book of Oromo? Do you think the developed primary text books and broadcasted news are written in line with this standard book?

10. H11: Kitaabni caaslugaa amma akka wabititti tajaajilaa jiruu kitaaba caaslugaa kan biroo adaafi turizimiitin maxxanfame. Kitaabni akuma durattii ibsametti matumasaatiin rakkinaafi hanqina waaltinaa qaba. Nutis yeroo sadaffaaf sirreesinee, kan itti dabalamuu qabus daballe maxxansuuf karoora qabna. Waan waalta'ee saniiniis hanqinni ittifayyadamaa baay'inaan jira. Kitaaboonni barataa sadarkaa tokkoffaa Afaan Oromotin barreffamanifi oduuwwan birootkaastiidhaan tamsaa'aan (keesumaayyu EBC, OBNfi FBC) kan nuti waaltassineen fayyadamurratti rakkina guddaatu jira. Kitaaboonnis ta'ee oduuwaan barreffaman keessatti rakkinni waaltinaa hedduu jira. Kan furuf ammo waliin hojjachaa jirra.

P11: The grammar book which is now used as a reference book is the book that is compiled and published by culture and tourism bureau of Oromia. Still, this book has to be edited gain as there are still issues to be standardized and we have to add what is left and republish it again. Textbooks of primary books and texts written by the electronic broadcast media (particularly OBN, EBC and FBC) are not in line with what we codified and there is variation. We are working to minimize such variation.

11. Q: Waaltina Afaan Oromootiin walqabatee haayyuulee kitaaba barataa sadarkaa tokkoffaa qopheessanif akkasumas gaazeexeessiitootaafi qaama dhimmi ilaallatuf caaslugarratti leenjiiwwan, simpoziyamii fi marii waliin gaggeessitan yoo jiraatee otoo ibsitan?

R: Explain if you have made any training, discussion or symposium regarding standardization of the language with language expertise particularly with primary school textbook developers and media expertise?

12. H12: Namootafi qaama armaan olitti caqafaman haga arratti waliin hojjachaa turree garafulduraattis kana caalaa waltii hiqnee hojjana. Waltajjii qorannoo adda addaa yeroo qopheessinu isaanis akka argamanii waa'ee waaltina afaanichaarratti yaada kennan godhaa turre. Leenjii walqabatee garuu hin keenninee.

P12: We have not made any standardization related training; but we have organized a symposium and a discussion regarding standardization of the language; particularly with primary textbook developers, media expertise and other concerned bodies. We are working together and we invite them when we organize research conferences.

13.Q: Ittii fayyadama xiinlatiin walqabatee rakkooleen waaltinaa mul'atan maal akka fakkaatuu otoo ibsitani?

R: Discuss if there are morphological and other variations related challenges which need to be standardized?

14. H7: Akka koree waaltina afaanititti jechoonis ta'ee latileen afuura tokkoon dubbifaman waltii hiqanii akka jecha tokkootittii haabarreffaman jenna. Kan jecha duraafi booda dhufanii jechootattii maxananii haabarreffaman jenna. Akkuma kanaan dura ibsiinetti haala qubeessutiin walqabatee nutiis 'guideline' yookin 'manual' waan qopheesinee maxxansinee ummataaf hin raabsineef rakkoolee amma argaa jirru kunmul'achaattii jiru. Kuun latilee maxxanseet barreessa kuun kophaa barreessa.

P7: From our point of view, those words and morphemes that will be read as one word should be written as one. Those affixes have to be written attached to the main word. One of the reasons for the existence of such variation is due to that we have not yet prepared and distributed guideline or manual to all concerned bodies and individuals.

15. H8: Afaan Oromoon haala ittifayyadama latilee amala garaagaraa qabu. Kan kophaasaani barreeffamu qabanis jiru. Kana otoo hinbarin namoonni adda addaa haala adda addaatiin kan barreessan jiru.

P8: Oromo morphemes exhibit different characteristics; some of them are written attached and some of them will be written alone. Various individuals are writing the same morpheme differently and it has to be normalized.

16. Q: Caasluga Afaan Oromoo kana caalaa waaltessuuf wanta hojjatamu qaburratti yaada maal qabdu?

R: What do you suggest for further standardization of Oromo?

17. H9: Akkuman sila jedhe duraa caasaa mana kana ol guddisu qabna. Hojiin afaan waaltessu waggaa baay'ee fudhata. Haayyulee bebbeekkamoo ta'anis ofittii hawwachuus qabna. motummaanis maallaqa ga'aa ta'eefi kan nu hojjisisuu danda'u ramadamuu qaba.

P 9: Like it has been mentioned earlier, first we have to improve the structure of the body from committee to language academy or higher status. Standardizing a language takes several years. Hence, we have to invite well experienced expertise. The government has to assign the necessary financial requirement that can assist to achieve our objective.

18. H10: Haga amma kan koreen hojjachaa jirruu isa 'descriptive approach' jedhamuuni malee isa 'prescriptive approach' jedhamunii miti. Afaan Oromoo kana caalaa waaltessuu kan dandeenyu isa 'prescriptive approach' jedhamun yoo hojjannee. Kana yoo ta'ee to'achuus ta'ee afaanicha kana caalaa kan waaltessuu dandeenyu.

P10: We have been standardizing the language through a descriptive approach and not a prescriptive approach. In order to standardize the language better we have to follow the prescriptive approach as it prescribes the usage of the normalized one.

III) Transcripts of focus group discussion three (FGD 3)

1. Q:- Marii kanarratti hirmaachuuf fedhii keessan waan argisiistaniif galatoomaa. Gosa barnoota kamiin akka eebifamtaniifi amma hojii kamirra akka hojjatan oto naaf ibsitani?

R: Thank you for your willingness to participate in the focus group discussion session. Would you tell me your field of graduation and occupation?

2. H1:- Tole: Maari nama hawwatuu kanarrattii argamee akkan hirmadhu waan na affeertef galatoomii. Ani Afaan Oromo maastarsiin qaba. Ani biiroo barnoota Oromiyaa keessatti ogeessa Afaan Oromotii. Kitaabonni barataa sadarkaa tokkoffaafi lammaffaa yeero qopha'aniis hirmadheera.

p1: Okay. I would like to thank you for inviting me to this interesting topic of discussion. I graduated with Afaan Oromo and I do have a Master degree in Afaan Oromo. I am serving as an expert of the language at Oromia education bureau. I also participated during textbook preparation

3. H2:- Ani maarii kanarratti akkan hirmaadhu waan naaffeerteef ulfaadhu jechuun barbaada. Ani 'educational planning and leadership'dhaan maastarsii hojjadheera. Goosa barumsuma kanaan biiroo barnoota Oromiyaa keessatti Ogeessa. Kitaaboonni sadarkaa tokkooffaafi lammaffaa yeero qopha'an hirmaadheera.

p2: I would also want to say thank you for having me on this discussion: I graduated with MA in educational planning and leadership. I am educational planning and leadership expert at Oromia education bureau and I also participated during textbook preparation.

4. H3:- marii kanarrattii waan naaffeerteef baay'ee galatoomii. Maastarii koo gosa barnootaa keemisriidhaan hojjadheera. Ammamoo biroo barnoota Oromiyaatti ogeessa saayinsii ummamaati.

p3: Thank you so much for inviting me to the discussion. I do have MA in chemistry and I am a natural science expert at Oromia education Bureau

6. H4:- Tole. Waan na afferteef galatomi. Ani barnoota lammummaafi amala gaariitii diigrii jalqabaatiin ebiffameera. Gosuma barumsa kanaatiin ogeessa

p4: Okay. Thank you for your invitation. I graduated BA with civics and ethical education and I am an expert in the area.

7. H5: -Tole! Anis waan na afferteef galatomi jechuun barbaada. Afaan Oromotiin diigrii lammaffaa hojjadheera. Anis goosuma barumsa kanaan ogeessa.

p5: Well. I also want to thank you for inviting me to the discussion. I graduated with MA in AfaanOromo and I am an expert in the subject.

8. H6:- Misha. Kun mataduree mariif tolu. Ani herregaan maastarsii qaba. Yeroo kitaaboonni sadarkaa tokkooffaafi lammaaffaa qopha'an hirmadheen ture; ammas Biroo barnoota Oromiyaa keessattii ogeessa herregaati.

p6: Good! It is a good topic to discuss. I am MA graduate of Mathematics and I also participated during textbooks. I am now a mathematics expert at Oromia education bureau

9. Q: Qubee Afaan Oromoo amma hojirra jiruu kana akkamittii ilaaltu?

R: How do you evaluate the current working Roman based script (Qubee) of Oromo?

10. H6:- Tole: Qubeen Afaan Oromo Afaan Oromo barreesuf kamiyyu caalattii mijataadha. Haata'u malee yeroo tokko tartiibnii qubee Afaan Oromoo A, B, C, D...irraa gara L, A, G, M... barumsaaf akka mijatutti jedhamee akka jijjiramee quban qaba. Tartiiba qubee jijjiruf h Tartiibnii qubee jijjiirra kun ammoo siyaasaan xaxaamee otoo mariifi yaada ogeessootaa hinfudhatii gara tartiiba qubee duraanittii deebi'era.

p6: Okay, Qubee Afaan Oromo is a script that best suit to write Oromo. But I remember some years ago, the alphabet order was changed to L, A, G, M... from A, B, C, D... for the sec of its suitability to teach. This reform did not include linguists and lexicographers. It was automatically politicized and reversed to A, B, C, D...without considering why education experts' opinion.

11.Q: Yeroo tartiibnii qubee jijjirammu sana biroon barnoota Oromiyaafi kareen waaltina Afaan Oromo haariroo qabuu (waliin hojjachaa turanii)?

R: Was there communication between the standardization committee of Oromo and Oromia education bureau during the reforms of the order of Oromo alphabet?

8. H2:- Gonkumaa. Hariiroo hinjiiru; waliinis matii tartiiba qubee kana kan foyyeessan. Biroon barnoota Oromiyaa kobumasaa jijjiree. Yaada rimeen jijjiiramichaas qubeen yeroo baay'ee deddebi'ee barreefamoota adda addaa kanaan dura barreeffaman kessaa argamee dura hiriruu qabaa kan jedhudha; kan hafanis akkasumatti. Jijjiirraa tartiiba qubee Afaan Oromorrattiis ta'ee haala qubeessusaarratti silaa kan shoora olaana tabachuu qabuu koree waaltina Afaan Oromoo ture.

p2: No! There was no communication between the committee and education bureau of Oromo. It is only education bureau of Oromo which reformed the reorder the alphabet. The bureau simple reordered the alphabet. The assumption was the order of the alphabet should be based on the frequency of each of the characters in various texts of Oromo. The character with highest frequency appears first in the alphabet and so on. It was the committee who has to play a great role in tasks of the reform of both the order of the alphabet and the orthography of Oromo in general.

9. H1:- Waan gaaridha kaaftee. Qubeen Afaan Oromo Afaan Oromo barreessuuf kan filatamee bara 1991 akka bara lakkoofsa alatiittii. Anii akkan yaadutti kanaan boodas qubee biraatiin waan jijjirramu natti hinfaakkatu. Qubeen Afaan Oromoo garuu ammas ilaallamuu qaba. Fakkeenyaaf, qubeewwaan cimdiin qubee tokkee booda otoo hiriranii barnootaf tola jidha. Kan jechuunis qubeen cimdiin dhumarrattii kaa'amurraa otoo A, B, C, Ch, D, Dh... N, Ny,...S, Sh,...P, Ph,...Z ta'anii tartiibaman wayya. Akkaataan qubeewwan cimdiin itti bocamanii nama qubee Afaan Oromo baratuus qubee cimdiin addaan basee baruuf yeroo fudhata jedheen yaada. Ch, Dh, Sh fi Ph'n haala walfakkaatuun bocaaman; dhumasaanirrattii 'h' qabu. nama yeroo duraatiif qubee Afaan Oromo baruuf qubeewwaan cimdiin afran kana addan baasanii baruun nama rakkissuu danda'a ta'a. Ny'n ammoo 'h' hin qabu. Gama biraatiin ammo, bareeffama Afaan Oromo keessattii yeroo baay'ee irra daddeebi'ee kan argamu sagaleen hudhaa (') haga ammaattii qubeen hin bocanneef. Kitaaboota barataa keessatiis ta'ee barreeffamoota biroo keessattii mallattoo waraabbii qeenxeedhaan barreeffamaa jira. Kun ammoo mallattoo tuqaalee waliin walfaalleessa. Kaanufuu qubeen bocamuufi qaba.

p1: Well, that is really a good question! Qubee Afaan Oromo was opted to be the official script of Oromo since 1991 and I think there is no way to replace it with other scripts. The script, however, needs reconsideration. For instance, pedagogically it is better to put the digraphs immediately after their respective monographs as in <A, B, C, Ch, D, Dh...N, Ny, ...S, Sh, P, Ph.. Z> rather than putting the digraphs at the end of the alphabets. Moreover, the formation of digraphs is confusing as four of them (<ch>, <dh>, <sh> and <ph>) are designed the same patten and <Ny> with different. It will be challenging for the beginners to differentiate the four digraph one from the other as four of them have <h> at their end. On, the other hand, frequently

occurring sound in the text of Oromo, the glottal stop is not yet represented with grapheme; it is being represented by the diacritic symbol in the textbooks which is also confused with punctuation marks. Hence, it has to be represented with a grapheme.

10. H5: Anis waan obboleewwan kiyya jedhanirrattii dabaluun barbaada. Sagaleewwan jechoota afaan birorraa gara Afaan Oromooti galaniif kan akka P, V fi Z fakkaataniif qubeen bocameefi otoo jiruu sagalee jechoota afaan alaa wajiin kan dhufee /z/'f qubeen waan hin bocamneef qubee adda addaattiin (kan akka Zh, Zy, Zn fi J) kitaaboota barataa keessaattii barreeffamaa jira. Kanaafu kuniis waalta'u qaba.

p5: I want to also add on what previously mentioned by my friends. Graphemes are design for the sounds that come with foreign words like p, v, z; on the other hand, the sound /z/, though not found in Oromo words, occurs frequently in the textbooks and is written in various graphemes such as <zh>, <zy>, <zn> and <j>. This is not yet codified and continues to be confusing to the lower grade learners. Hence, it has to be codified.

11. Q: Afaan Oromoon kitaaba barataa sadarkaa tokkoffaafi lammaffaa keessatti akkamittii gabbataa jira?

R: How is Oromo elaborated in the textbook of primary and secondary textbooks?

12. H3:- Jechoota biyyaa alatiin yaadni alaa akka ibsamaa akka jiran kitaaboota barataa sadarkaa tokkooffaafi lammaffaa saayinsii ummamaa ni argina. Jechoota liqeeffachuudhan alas, yaadni adda addaa alaas Afaanuma Oromotiin kitaaboota kuneen kana keessattii yeroo ibsamantu jiruu. Yeroo tokko tokkos hikni yaada afaan biroo gara Afaan Oromooti hikkamaas jiruu. Yaadni lamaafi lamaa ol ta'anis yeroo tokkoo tokkoo hiika jechaa bal'isuu fi dhiphisuudhaan yeroon itti ibsamus jira. Jechoonni lamaafi isaa ol gabaajeedhaan (abriveeshinidhaan) fi jechoota tishoodhaan yeroon itti barraa'anis jiru.

P3: I have seen foreign concepts being included in the natural science textbooks via borrowing. In some of the cases, some terms are also seen to be expressed by the native equivalent. In some

cases, non-Oromo concepts are translated in Oromo. In some other cases, the two or more concepts are expressed using single form through extending meaning. There are also other means such as an abbreviation and compounding.

13. H4:- jechooni Afaan Oromo afaan birarraa liqeeffatee (liqii kalattii kan akka 'dimookraasii' fi alkalattii), hiikee, jechooni hiika dabalataa akka qabaatan godhee afaanichii akka gabbattuu ta'aa jira.

p4: Several techniques such as borrowing, translation, meaning extension have been used to elaborate the language. There are direct borrowings such as 'democracy' and so on

14. Q:-waaltina Afaan Oromoo walqabatee, hudhaawwan kitaaboota sadarkaa tokkooffaafi lammaffaa kessattii mul'atan maal fakkaatu?

R: What are the existing challenges of standardization of Oromo in the primary and secondary school textbooks?

15. H2:-Misha, imala Afaan Oromo waaltessuu keessatti gufuwwan adda addaatu kitaabota barataa keessattii jira. Jechooni yeroo tokko walitti maxananii yeroo kaan ammoo kobakobaasaanii yeroo biroo ammoo jechoota giduutti sara kaa'uudhaan kitaabootuma kanneen keessattii barreeffamaniiru. Fakkeenyaaf, *fi* 'and' fi *hin* yeroo tokko irraa hiqanii, yeroo biraa itti maxanani barreeffamaa jiru. Fakkeenyaaf *anaafi inni*, *anaa fi inni*; *hindeemnee hin deemne* fi kan kana fakkaatan.

p2: Well, there are several challenges in the textbooks that became an obstacle to the standardization process of the language. Words are sometimes written together in another time apart, sometimes with a hyphen, and in other times without hyphen as in, for instance, *anaafi inni* and *anaa fi inni* 'I and him', *hindeemnee* 'I have gone' and *hin deemne*, attaching and detaching the negation marker from the word that follow and so on.

16. H1:- Rakkowwan afaanich waaltessuu keessattii mul'atan baay'eedha. Yaadni tokko qooftii looga adda addattifi jechoota adda addaatiin barreeffamaa jiruu. Fakkeenyaaf, *soba* jedhu haala

saddin barreeffamaa jira; *soba, dhara, kijiba*; *jechii afaan Ingliizitiin* 'Tuesday' jedhus yoo xiqaaatee haala shaniini *Baakaa boodaa, Facaasaa, Kibxata, Lammaffo, Salaasaa* jedhame barreeffamaa jira. Garaagarummaa looga Afaan Oromo keessa jirun haalli itti fayyadama jechoota rooga dhabeera. Jechoota gabaabsanii (abriveet godhanii) barreessuuf seerii qaama dhimmi ilaalatuun hinbaafnee. Waan seerii hinjirree abbaan barbaadefi barreesitooni akka fadhanitti gabaasani barreessaa jiruu. Abraveeshiin tokko haala lamaafi isaa oliin barreeffamaatii jira. Fakkeenyaaf, biroo barnoota Oromiyaa jedhani barreessuuf BBO, B.B.O fi yeroo tokko tokko ammo B/B/O godhaniitis kan barreessan jiru. Seerii jechoota jemakafi digalafi itti barra'anis waan tolchamneefi hinjireef namu akka fadhetti barreessaa jira.

p1: There are many existing challenges of standardization of the language. There are many instances in which a single concept is expressed with several forms (words). For example the concept 'false' is expressed in three different ways *soba, dhara, kijiba*; 'Tuesday' in five ways as in *Baakaa boodaa, Fac'aasaa, Kibt'ata, Lammaffo, Salaasaa*. This is due to lexical variations in Afaan Oromo. There are no general rules for abbreviations in Afaan Oromo, and that is why every writer creates abbreviation in any form he /she wants. A single abbreviation is written in two or more forms as in *BBO, B.B.O* and even sometimes with slash *B/B/O*. There is also no codification as to how blends and compounds should be written. Hence, anyone is writing such words as sh/he wants.

17. H6:- Koree waaltina Afaan Oromo afaanicha waaltessuf seeran waan keennamef jechootafi kanneen biroo waaltessuufi moggaasudhaan garaagarummaa afaanicha keessaatti mul'atu dhiphisun isarraa eegama. Kan hin waaltoofnee kan waalta'erra baay'ee caala. Kan moggaafaman gaara hinjiraan jechisissutti nama geessa. Liqii jechoota alaa hin barbachifne liqeeffachuun to hin hanqatiin seeri jechooni liqeeffaman akkamitti barraa'u akka qabanillee seerri hinjiru.

p6: The Standardization committee of Oromo is a legitimate body of Oromia to work on codification to minimize the existing variations. But, what is codified is almost none as compared to what is to be codified. There are no rules as to how borrowed words should be written. There is also unnecessary borrowing.

18. H5:- Kan waalteffamu qaban Afaan Oromo keessatti garaagarummaa hedduu jira. Fakkeenyaaf, Loogootni Afaan Oromoo lakkoofsa sadarkaa haala adda addaatiin barraa'a jira. Fakkeenyaaf, isa 'first, 1st ' jedhuuf tokkoffaa., tokkeesso, tokkossa, tokkeysoo, waan kana akkaatanin barreeffamaa jira. Latilee danooma agarsiisan *-oota, -ota, -eenfi* kan kana fakkatanin yeroo tokkoo san yeroo biraa ammo kaani lamaafi lamaa oliin bakka busnee fayyadamaatii jirra. Fakkeenyaaf, isa afaan Ingiliziitiin 'horses' jedhuf haala lamaan, fardootafi fardeen gonee barresina. Kitaaboota keessattis kan agarru kanuma. Garaagarumman kanbiroo akkasitis jiru. kuneen ani kaasee akka fakeenyaa haata'u jedheen.

p5: There are a number of variations in Oromo that has to be standardized. For instance, there are variations in expression of Oromo ordinals as in, for instance *tokkoffoo, tokkeesso, tokkossa tokkeeysoo, tokkooffaa* and so on. Plural forms words are also seen to be represented in more than two ways. For example 'houses' seen to be marked both with *-oota* as in *manoota* and *-een manneen*; 'horses' as *fardoota & fardeen* and so on. There are several such variations; I only state this as an example

19. Q:- Waaltina Afaan Oromootiin walqabatee, hariiroo dhimmamtootaa akkamitti ilaaltu?

R: How do you evaluate the coordination among stakeholders' to standardize Oromo?

20. H3:- Koreen waaltina Afaan Oromo jechoota moggaasaa jira. Dhaabni kun baay'e duubattii harkifataadha. Waan waalteffamuu qaban hin jalqabamne jechuu dandeenya. Kun kanaan otoo jiru, kan waalta'anirree hojirra olaa hinjiran. Dhimmamtooni waaltina afaanii kan akka koree waaltinaa, biroo barnoota, barreesitoota, midiiyaa, mitimootummaafi dhaaboota motummaa hariiroo hinqaban jechuu dandeenya. Yaadni saayinsii ummamaafi saayinsii haawaasaa kan ibsaman bareesitoota garaagaratiin haala garaagaraatinii.

p3:- The standardization committee of Oromo is codifying words. But, the process is very steady. The issues to be standardized are almost untouched. Moreover, what are codified are not put into practice. The coordination between stakeholders such as the committee, the education bureau, writers, media, and various governmental and nongovernmental institutions is almost

none. The scientific and social science related concepts are designated by various individuals in various ways.

21. H4:- Hariroon waaltina Afaan Oromoofi dhimmamtoonni kanaa olittii tuqaman kan biroo cimaa ta'uu qaba. Waantii waalteeffamanifi moggaafaman wiirtuu tokkorraa tamsa'ee haala walfakkaatuun dhimmamtoota hundaan ittifayyadamuu qabu.

p4:- There must be a strong relation between the standardization committee of Oromo and other stakeholders mentioned right now. The codified and standardized lexical items should be disseminated from only one center so as to be used uniformly by all stakeholders.

22. H1:- Itti fayyadama jechootaa kitaaboota barataa sadarkaa tokkoffaafi lammaffaa keessatti garaagarumman mul'atan baay'ee bal'aa. Koreen waaltina Afaan Oromoo kitaaboonni barataa yeroo qopha'an hin hirmaannee. Otoo hirmaatanii garaagarummaan akkasii kiboota keessattii hin mul'atu ture. Waan hariroonfi hojiirraa olmaan waaltteeffamaniifi moggaaffamanii hinjirreef garaagarummaan affanichii keessaatti mul'atuu daran baal'achaa jira. Hariroon dhimmamtootaa waan garmalee waan hin mul'anneef waaltina afaanichaarrattii gufuu ta'aa jira.

p1: There are several lexical variations in the primary and secondary textbooks. The standardization committee of Oromo did not participate during the preparation textbook. Had the committee participated during the preparation, there couldn't have been such a variation in the textbooks. The absence of communication and not implementing the codified lexical and other various rules is now destandardizing the language. Hence, this loosen is affecting the standardization of the language.

23. Q:- Yaada Afaan Oromorraa adda ta'aan akkamittii ibsitu (moggaaftuu)?

R: How do foreign concepts are included in the textbooks?

24. H1:- Yaadni afaanichaa ala jiru haala adda addaatiin ibsama. Yaada ibsuuf jechootni afaanich keessa otoo jiranu jechoonni afaan biraarra yeroon liqeenfamanii yaadicha ibsan kitaaboota

keessatti argineerraa. Yeroo biraa ammoo jechaa liqeeffaannee jecha tokkon qofa ibsuu otoo danda'amu jechoota Afaan Oromoo sadii oliin (otoo qusannaa iddo yaada keessa hingalchiin) kitaaboota barataa keessatti ibsamaniru.

P1: Foreign concepts are expressed in various ways. I have seen in the textbook that despite the existence of genuine words that may represent the word, a concept is seen to be represented by borrowing words. In some other cases, a foreign concept which can be represented with a single word via borrowing, the concept is seen to be represented with more than three native words of the language, which is not economical as compared to a single borrowed word.

25. H6:- Kitaabonni barataa, keesumaayyu kitaabonni barata saayinsii ummamaa yaada afaan Oromootiin alaan (foreign concepts) kan gutamanidha. Dhimmamtoonni waaltina afaanicha otoo waliin hojjatan ta'ee yaadin tokko tokko jechoota otoo hin liqeeffatiin jechootuma Afaan Oromotiin ibsamuu danda'u turan. Amman ani beekuttii qaamnii seeran beekamu jechoota saayinsii moggaasee haala walfakkaataa ta'een akka dhimmamtoonni ittiffayyadaman kan gochaa jitu hin jiru. Garaagarummaa amma kitaaboota barataafi iddoowan adda addaatti sababni nutii argaa jiruu kanuma

p6: Textbooks of Oromo, particularly textbooks of natural science are filled with several foreign concepts. Had the stakeholders of standardization of Oromo worked together some of the foreign concepts would have been expressed by native words. To the best of my knowledge, there is no legitimate body which is working to codify the scientific concept to be uniformly used by stakeholders. That is the basic reason for the existence of lexical variation in the textbooks and other texts.

26. Q:-Afaan Oromoo waaltessuuf waaltu hojjatamuu qaba jittu?

R: What do you think needs to be done to standardize Oromo?

27. H1:-Gaaridha, tokkooffaan, caasaa waaltina Afaan Oromoo koreerraa gara sadarkaa itti anutti ol guddachuu qaba; gara akkadamititti ol guddachuu qaba jechudha. Ogeesoonni buleeyyin kan akka piroffesaroota gosa qorannaa xiqooqaa, jechootaa, barumsaafi gosoonni qorrannoo

biroos waaltinaa afaanichaa keessattii hirmaachuu qabu. Kana ta'uu baannaan, na amanii afaanichii waalta'uurraa (walfakkaachurra) garaagarummaan baay'ee ballachaatti deema. Lammaffaa, waaltina afaanichaatiin waalqabatee dhimmaamtonnii hundii waan waalta'anfi mooggaaffaman haga danda'ameetti haala waalffakkaatuun itti faayyadamuf hariiroosaani cimsuu qabuu.

p1: Good! First, the status of the standardization committee of Oromo has to move from the status of committee to the next higher status i.e. it must be lifted to language academy. There must be also high level professionals such as professors of linguistics, lexicographers, and education and necessary areas. Otherwise, believe me; the language will be more varied than becoming more uniform. Second, other stakeholders have to have a strong bond with the standardization committee of the language so as to use the codified items as uniformly as possible.

28. H6:-Tole, waan waalta'an dhimmamtootaf waan hingeenyeeff, koreen waaltinaafi dhimmamtoonni waaltinaa waan walin hojjachaa hin jireef waan waalta'eeffi moggaafamee hojiirra olaa hin jiru. Kitaaboonni yeroo qophaa'an hirmaanaan korichaa dabalatee, ogeeyyiin xinqooqaa, ogeeyyii barnoota, ogeeyyii jechoota waliin hojjachuu qabu. Barreeffamnii kitaaba barataa qopheessuf olu kamiyyuu kan waalta'ee ta'uu kan qabuu barattoonnii kan barachuu qaban isa waalte'ee sana waan ta'uu qabuufi. Kanaan olitti waan faallaa ta'een kitaaboonni amma sadarkaa tokkoffaafi lammaffaa amma kan ittin qopha'an afaan waalta'eenii miti. kitaaboota keessa garaa garummaa adda addaatu mal'ata. Biroon barnoota Oromiyaafi koreen waaltina afaanichaa waliin hojjachaa hinjiran. Lamanuu gamaa gamana taa'anii hojjachaa jiruu. Ogeeyyiin kitaaba barataa sadarkaa tokkoffaafi lammaffaa qopheessaniis waan waaltinni afaanichaa waaltessee qopheesse dhisanii jechoota ofiif bocachaa jiruu.nutii kopha kophaa hojjachaa jirraa jechuun ni danda'ama.kun ammoo afaanichi akka hinwaaltofnee godha.

p6: Okay, what is codified is not put into practice due to lack of distribution of the codified items and lack of coordination among the standardization committee of Oromo and other stakeholders. Textbook preparation should involve standardization personnel, linguist, educational expert and lexicographers. The language to be used in any textbook shall be a standard one as students are

expected to learn standard language. As opposed to the above, textbooks are not prepared with the standard Oromo and there are several lexical and orthographic variations. The education bureau of Oromia and the standardization committee of Oromo is not working together. Textbook writers elaborate the language via forming new terms to designate scientific and technical terms rather than using what has been codified by the standardization of Oromo.

IV) Transcript of interview with head of center for culture of Oromia

1. Q:- Koree waaltina Afaan Oromoo yoom hundaa'e? Ga'een isaa maali?

R: When was Oromo standardization committee established? What is its role?

2. H:- Koreen waaltina Afaan Oromoo bara 1985 A.L.E dhumarra hundaa'e. Ge'ensaa waa'ee dhimma afaanicha waaltessuurratti hojjachuudha. Afaan Oromoo karaa saayinsaawaa ta'een hojjachudhaan afaanicha afaan saayinsiifi teeknoolojii akka ta'u, ummanni caasaa, jechootafi haala qubeessuu walfakaatuun akka barreessuu danda'u gochuudha.

P:- Standardization committee of Oromo has been established in 1991 and its responsibility is standardizing the language. It has to study the language scientifically to make language of science and technology. In addition to this, the committee has to work to make the language maximal variation in function and minimal variation in form and thereby influencing the speaker to write the language with minimal variation.

3.Q:- Koreen kun waaltina afaanichaatiin walqabatee haga ammaatti haayyulee adda addaa (fakkeenyaaf haayyulee barnootaa, miidiiyaa, barreessitootafi kkf) waliin mari'atee beekaa?

R:- Explain if the committee has ever made a discussion with language expertise, media expertise, authors?

4. H:- Nuti miseensoota kanaa olitti caqafaman waltajii keenya adda addaarratti akka hirmaatan goneerra, kanaan duras hirmaatanii beeku. Qoranno yoo gaggeessinu isaanis akka miseensatti ni afferra haarcaatiin akkuma jirutti ta'ee. Korarrattii mariinis ni godhama.

P:- We will invite all those individuals you mentioned earlier on our various discussion and research conferences and they indeed attended the discussions and conferences. They have participated on previous discussions as they are members of standardization councils of Oromo.

5. Q:- Kan haga ammaattii waalteessitan akkamittii ummata biraan geessu?

R:- How do you distribute the codified norms to the users?

6. H:- Maxxansinee wiirtudhaan ummataaf raabsaa jirra. Yeroo baay'ees ta'u baatu yeroo tokko tokko midiyaadhanis jechoota waalteffaman ummata ga'aa jirra. Mariinis tamsaasa kallattidhan yeroo tokko tokko ni taasifama.

P:- We codify and print it on wiirtuu 'center' and then we distribute it to the concerned bodies. We also sometimes distribute the codified lexicon to some selected media and they will prepare a program to discuss about the codified norms and words with viewers (speakers).

7. Q:- Afaan Oromoo waaltessuun walqabatee yaada ummata kanaan dura qoratani yoo beektan ta'e otoo ibsitani?

R:- Have you ever done any assessments (through oral discussion or written) regarding the reaction of Oromo speakers towards standardization of the language?

8. H:- Eeyye qorannee beekna. Ummaniis afaanichii akka waalta'u akka barbaadan barreerra. Kana waan ta'eef waggootan digdamii shaniif afaanicha waalteessaa jirra. Kan haga amma waalteessine qorachuuf ummata keessaa bakka buutota ummataiddoo barbaachisaa ta'ee hundarraa filannee haala hirmachisaa ta'een yaada fudhachaa jirra. Wanta koreen kun hojjaterree haayyoonni qoratani ummataaf dhiyeessaa jiru. Wiirtuurrattiis akka maxxansan gooneerra.

P:- Yes. We have conducted a study in order to investigate the necessity of standardizing the language. We realized that there is a need of standardizing the language from the speaker side and that is why we have been working on standardizing the language for the last 25 years. Still, we selected various individuals from various spheres and of the varieties that represent the

speakers of the language in order to evaluate what we have codified. In addition to our research various researchers conducted a research related to standardization of the language and they also presented it in our research conference.

9. Q: Afaan Oromoo haga ammaa waalteessitan fudhatama argachuufi dhabuu isaanii akkamittii mirkaneeffattu?

R: How do you check whether the codified norms have got acceptance (being implemented) or not by language users?

10. H. Toftaa adda addaatin mirkanessina. Korarratti afaanicha haga ammaattii hagam waalta'ee? Wanti sirraa'u malu maali? kan haga ammaatti moggaafnee hagam fudhatamaa jira? jennee bakka buutoota afaanicha dubbatanifi iddoo adda addaarra dhufan gaafanna. Qorannoos gaggeessinee beekna. Argannoon qorannoo keenyaa, kanneen haayyuleen biroon qorataniifi yaanni korarrattii ka'u kan mul'isuu hanqinni jiraatuus kan haga ammaattii moggaasaafi waaltessaa jirru ummataa fudhatamaa akka jiruudha.

P: We use various techniques in order to check whether the codified norms have got the acceptance (being implemented) or not by language users. First, on the conference, we organize; we ask speakers' representative who are from various spheres and varieties of the language to discuss to what extent the codified norms are being accepted. The results of the study we conducted and the discussion on the various conferences show that what we codify has got an acceptance though there are still gaps.

11. Q: Gama motummaatiin Afaan Oromoo afaan hojii motummaa feedaralaa godhuuf karoorri jiraa? Yoo karoorri jira ta'e, afaanicha kana caalaa waalteessuf tattaaffin kana caalaa godhamuu maali?

R: Is there any plan to elevate the language into the federal official working language? If it is to be elevated to the next higher official status, what extra efforts has to be made in standardizing the language?

12. H: Afaan Oromoonafaanoota Afrikaa keessaa baay'ina dubbattoota afaanitiin sadarkaa sadaffaarrattii argama. Afaan tokko afaan hojii fedaraalaa godhuuf ulaagaalee fudhataman keessaa baay'ina dubbattoota afaanichaa 'numerical status', afaanichi tokkummaa biyyaatiif gaheen inni qabu, gama dinagdeetifi siyaasa biyyattiif faayedaan afaanichi qabuu maali jedhamee ilaallamu qaba. Ani gama kootin Afaan Oromoon ulaagaa afaan tokko afaan feedaraalaa ta'uuf gaafatuu niguutan jedha. Motummaan naannoo Oromiyaatis gaaffii kana motummaa feedaraala gaafachaa jira. Karoorriis jira jenne fudhachuu dandeenya. Kun ammo dhihootti furmaata ni argata. Afaanichi sadarkaa itti anuttii yoo guddatee ammo afaanicha keessa adda addummaan haala qubeessuu, jechootaa, latiwwanifi kan kana fakkataan baay'inaan wanta jiraniif afaanichi kana caalaa waalteffamuu qaba.

Akka qorannoodhaan mirkaneessine akka mul'isuu yoo ta'e waaltinni afaani jidugala tokkorra tamsa'uu qaba. Waaltinni afaanii bakka adda addaa gaggeeffama taanaan waaltessuu odoo hinta'in gargar fageessuu ta'a jechuudha. Waaltinni Afaan Oromoo amma sadarkaa koreerra jira. Kana gara caasaa akkaadaamitiittii guddisuu qabna. Koreen waaltinaa haga ammatti maleenya ibsaa 'descriptive approach' hordofaa ture. Kana keessattii filannoodhuma dhiyeessita. Kana kana moggaafneerra (waalteessineerra). Otoo kana kana hojirra oo'lchitanii jenna. Garuu yoo waan waaltesinee kana hojirra oolchuu baatan dhibbaa kan akkanaafi kan akkasii fida hinjennu. Kun ammo amala rakkina maleenya ibsaati. Karaa biraatiin ammo waanin to'annuufi adabnun lafa hingoone. Kana booda waaltinni afaanichaa kana caalaa akka hojjatamu gara maleenya ajajaattitii 'prescriptive approach' darbuu qaba. Koree kana ammo saddarkaa olaanaatitti haangessuu qabna.

P: Oromo is the third most spoken in Africa with regard to the number of speakers. To elevate any language to the next higher official status (federal official working language), at least the following criterion has to be considered. The first is the number of speakers of that language. The second is considering the role that language play in unifying the people of that country. The third is the political advantage of that language. From my point of view, Oromo fulfills the criterion needed to be the official working language of the federal government of the country. The Oromia regional government has also requested the federal government to reconsider the federal official working language of the country. Hence, we can consider that there is a plan to

elevate Oromo to the federal official working language of the country. I think, this will be resolved in the near future. We have to work more to standardize the language as there are numerous orthographic, lexical and other standardization related issues to be considered.

Various studies show that there should be only one central institution that works on standardization of the language. What has been codified must also be distributed to the speakers of the language from the same center. To the contrary, if language continues to be codified by various individuals here and there, the difference will be wider and wider rather than being minimized. Tasks of standardizing the language are being done at the level of the committee. But, from the past month on wards, the next higher body called the Oromo standardization council was established. Still, it has to be elevated to the language academy level. The committee is working for the last 25 year using descriptive approach and not a prescriptive. When we follow the descriptive approach, it means that we have codified this and that; it will be good if you use this and that. In this case, we only present the existing options. We did not put any checking, controlling and correcting mechanism for those who are using or codifying against what we have codified as a legal institution. This is due to the nature of descriptive approach we followed. But now on wards, prescriptive approach has to be applied in order to standardize the language better than what we did. Like I have stated earlier the status of the institution has to also be elevated to the language academy so as to make the standardization process more effective.

Appendix H: List of texts of the corpus

No	Subject	Grade (Volume)	Year of publication	Publisher
1	Afaan Oromo (Oromo)	Grade 1	2005 E.C /2012 G.C	EBO
2	Afaan Oromo (Oromo)	Grade 2	2005 E.C /2012 G.C	EBO
3	Afaan Oromo (Oromo)	Grade3	2005 E.C /2012 G.C	EBO
4	Afaan Oromo (Oromo)	Grade 4	2005 E.C /2012 G.C	EBO
5	Afaan Oromo (Oromo)	Grade 5	2005 E.C /2012 G.C	EBO
6	Afaan Oromo (Oromo)	Grade 6	2007E.C/2014 G.C	EBO & MoE
7	Afaan Oromo (Oromo)	Grade 7	2007 E.C/2014 G.C	EBO & MoE
8	Afaan Oromo (Oromo)	Grade 8	2007 E.C/2014 G.C	EBO & MoE
9	Afaan Oromo (Oromo)	Grade 9	2005 E.C/2012 G.C	MoE
10	Afaan Oromo (Oromo)	Grade10	2005 E.C/2012 G.C	MoE
11	Afaan Oromo (Oromo)	Grade 11	2005 E.C/2012 G.C	MoE
12	Afaan Oromo (Oromo)	Grade 12	2005 E.C/2012 G.C	MoE
13	Barnoota Herregaa (Mathematics)	Grade 1	2003 E.C /2011 G.C	MoE
14	Barnoota Herregaa (Mathematics)	Grade 2	2003 E.C /2011 G.C	MoE
15	Barnoota Herregaa (Mathematics)	Grade 3	2003 E.C /2011 G.C	MoE
16	Barnoota Herregaa (Mathematics)	Grade 4	2003 E.C /2011 G.C	MoE
17	Barnoota Herregaa (Mathematics)	Grade 5	2003 E.C /2011 G.C	MoE
18	Barnoota Herregaa (Mathematics)	Grade 6	2003 E.C /2011 G.C	MoE
19	Barnoota Herregaa (Mathematics)	Grade 7	2003 E.C /2011 G.C	MoE
20	Barnoota Herregaa (Mathematics)	Grade 8	2003 E.C /2011 G.C	MoE
21	Saayinsii Naannoo (Env. Sc.)	Grade 1	2005 E.C/2013 G.C	EBO
22	Saayinsii Naannoo (Env. Sc.)	Grade 2	2005 E.C/2013 G.C	EBO
23	Saayinsii Naannoo (Env. Sc.)	Grade 3	2005 E.C/2013 G.C	EBO
24	Saayinsii Naannoo (Env. Sc.)	Grade 4	2005 E.C/2013 G.C	EBO
25	Asteetiksii (Aesthetics)	Grade 1	2005 E.C/2013 G.C	EBO
26	Asteetiksii (Aesthetics)	Grade 2	2005 E.C/2013 G.C	EBO
27	Asteetiksii (Aesthetics)	Grade 3	2005 E.C/2013 G.C	EBO
28	Asteetiksii (Aesthetics)	Grade 4	2005 E.C/2013 G.C	EBO
29	Barnoota Hawaasaa(Social study)	Grade 5	2003 E.C /2011 G.C	MoE
30	Barnoota Hawaasaa(Social study)	Grade 6	2003 E.C /2011 G.C	MoE
31	Barnoota Hawaasaa(Social study)	Grade 7	2003 E.C /2011 G.C	MoE
32	Barnoota Hawaasaa(Social study)	Grade 8	2003 E.C /2011 G.C	MoE
33	Saayinsii Walsimataa(interim Science)	Grade 5	2004 E.C/ 2012 G.C	EBO
34	Saayinsii Walsimataa (Interim Science)	Grade 6	2004 E.C/ 2012 G.C	EBO
35	Barnoota Lammummaa fi Amala garaarii/Civics & Ethical education	Grade 5	2003 E.C /2011 G.C	MoE
36	Barnoota Lammummaa fi Amala garaarii /Civics & Ethical education	Grade 6	2003 E.C /2011 G.C	MoE
37	Barnoota Lammummaa fi Amala	Grade 7	2003 E.C /2011 G.C	MoE

	garaarii/Civics & Ethical education			
38	Barnoota Lammummaa fi Amala garaarii /Civics & Ethical education	Grade 8	2003 E.C /2011 G.C	MoE
39	Barnoota guddinaafi jabeenya qaamaa (Health and physical fitness education)	Grade 5	2003 E.C /2011 G.C	MoE
40	Barnoota guddinaafi jabeenya qaamaa (Health and physical fitness education)	Grade 6	2003 E.C /2011 G.C	MoE
41	Barnoota guddinaafi jabeenya qaamaa (Health and physical fitness education)	Grade 7	2003 E.C /2011 G.C	MoE
42	Barnoota guddinaafi jabeenya qaamaa (Health and physical fitness education)	Grade 8	2003 E.C /2011 G.C	MoE
43	Barnoota muziqaa (Music)	Grade 5	2003 E.C /2011 G.C	MoE
44	Barnoota muziqaa (Music)	Grade 6	2003 E.C /2011 G.C	MoE
45	Barnoota Ogummaa Gaarii/Fine arts)	Grade 5	2003 E.C /2011 G.C	MoE
46	Barnoota Ogummaa Gaarii/Fine arts)	Grade 6	2003 E.C /2011 G.C	MoE
47	Bayoolojii (Biology)	Grade 7	2006 E.C/2013 G.C	MoE
48	Bayoolojii (Biology)	Grade 8	2006 E.C/2013 G.C	MoE
49	Keemistiirii(Chemistry)	Grade 7	2006 E.C/2013 G.C	MoE
50	Keemistiirii(Chemistry)	Grade 8	2006 E.C/2013 G.C	MoE
51	Fiisiksii (Physics)	Grade 7	2006 E.C/2013 G.C	MoE
52	Fiisiksii (Physics)	Grade 8	2006 E.C/2013 G.C	MoE
53	Caasluga Afaan Oromo (Oromo Grammar	Volume 1	1998 E.C/2006 G.C	CTBO
54	Qajeelcha/Wiirtuu jildii 1-5/ (Reviezed Wiirtuu books volume 1-5)	Volume 1-5	2000E.C/2008G.C	CTBO
55	Wirtuu Jildii 1 /Wiirtuu volume 1	Volume 1	1987E.C	CTBO
56	Wirtuu Jildii 2 /Wiirtuu volume 2	Volume 2	1989 E.C	CTBO
57	Wirtuu Jildii 3 /Wiirtuu volume 3	Volume 3	1990 E.C	CTBO
58	Wirtuu Jildii 4 /Wiirtuu volume 4	Volume 4	1991 E.C	CTBO
59	Wirtuu Jildii 5 /Wiirtuu volume 5	Volume 5	1992 E.C	CTBO
60	Wirtuu Jildii 6 /Wiirtuu volume 6	Volume 6	1993 E.C	CTBO
61	Wirtuu Jildii 7 /Wiirtuu volume 7	Volume 7	1995 E.C	CTBO
62	Wirtuu Jildii 8/Wiirtuu volume 8	Volume 8	1999 E.C	CTBO
63	Wirtuu Jildii 9 /Wiirtuu volume 9	Volume 9	2001 E.C	CTBO
64	Wirtuu Jildii 10 /Wiirtuu volume 10	Volume 10	2005 E,C	CTBO
65	Wirtuu Jildii 11 /Wiirtuu volume 11	Volume 11	2011 E.C	CTBO

Appendix I: Summary of expertise of SCO who participated in FGD 1

Place of data collections: Addis Ababa

No	expert	age	sex	academic level	experience	position	date of FGD1
1	A	42	M	MA /Linguistics, PhD candidate	16	Researcher	May29, 2018
2	B	48	M	BA /Linguistics	24	Researcher	
3	C	28	M	BA/Afaan Oromo, MA candidate	6	Researcher	
4	D	32	F	BA/ Afaan Oromo	12	Researcher	
5	E	38	F	MA/ Afaan Oromo	14	Researcher	
6	F	32	M	BA/ Afaan Oromo	10	Researcher	

Appendix J: Summary of expertise of SCO who participated in FGD 2

Place of data collections: Addis Ababa

No	expert	age	sex	academic level	experience	position	date of FGD
7	G	34	M	BA/Afaan Oromo, MA candidate	14	Researcher	3 Oct 2019
8	H	32	F	BA/ Afaan Oromo	10	Researcher	
9	I	36	M	MA/ Afaan Oromo	12	Researcher	
10	J	35	F	BA/ Afaan Oromo,MA candidate	12	Researcher	
11	K	35	F	MA multilingual & Multicultura	15	Researcher	
12	L	36	F	BA/ Afaan Oromo, MA condidate	16	Researcher	

Appendix K: Summary of interviewee of head of center for culture of Oromia

No	Expert	age	sex	academic level	experience	position	date of FGD
1	M	43	M	MA/ Afaan Oromo, PhD Candidate	14	Researcher	31 May 2018

AppendixL: Summary of expertise of education who participated in FGD 3

Place of data collections: Addis Ababa

No	Expert	age	sex	academic level	experience	position	date of FGD3
1	N	45	M	MA/Afaan Oromo	18	Expert	8 Nov.2019
2	O	53		MA/educational planning and leadership	28	Expert	
3	P	37	F	MA/ Chemistry	19	Expert	
4	Q	34	F	BA civics & ethical education	13	Expert	
5	R	39	M	MA/ Afaan Oromo	20	Expert	
6	S	40	M	MA/ mathematics	19	Expert	

Appendix M: Short list of words codified by SCO

No.	Oromo	English	Amharic
1	abaloo	apostrophe	እንቶኔ
2	abbagaar	ombudsman	እንባ ጠባቂ
3	addabaasoo	marker	ማርክር
4	afaanee	oral	ቃላዊ
5	afarsa/qilleesisuu	aeration	አናፋሽ
6	afarsuu	to fan	ማንፈስ
7	afata	Carpet	ንጣፍ
8	afdhaba	language lose	የቋንቋ መሰገን
9	aflaka	language shift	የቋንቋ መቀየር
10	afmaka	code mixing	የቋንቋ መቀየጥ
11	aftara	code switching	የቋንቋ መቀላቀል

12	afuddaa	Diaphragm	ዲያፍረም
13	agarsiisa dandeettii addaa	talent show	የልዩ ችሎታ ማሳያ
14	agoja	demonstrations sight	ሰርቶ ማሳያ
15	akkaadaamii	Academy	አካዳሚ
16	akkaan, haalaan, rogaan	properly	በአግባቡ
17	akkaataa wabeeffannaa	citation style	የማጣቀሻ ስልት
18	akkamii	dimension	ልክ/ግጥ/ ስፍረት
19	akkee	smiley	ተነጻጻሪ ዘይቤ
20	alabsuu	Discriminate	መለየት
21	alalaasa	intone	ማንጎራጎር
22	albaabsummaa	Liberalism	ሊብራሊዝም
23	alboola	tropic of Capricorn	የትሮፒካል መስመር
24	aloola	fieldwork	የመስክ ስራ
25	alsaayinsaawaa	non scientific	ኢ ሳይንሳዊ
30	alxaa	appetite	አፒታይት
31	amala	characteristics	ባህሪይ
32	amala	conduct	ጸባይ
33	amanamaa	loyal	ታማኝ
34	amansa	adaptation	ስምምነት
35	amansaa	adapter	አስማሚ
36	amansiisuu	Indoctrination	ማሳመን
37	amansuu	adapt (to)	ማስማማት
38	anguu	Stanza	አንጻ
39	araara	to negotiate	መደራደር
40	araara	Reconciliation	ዕርቅ

41	aramuu/ shuqunuu	to weed	ማረም
42	arbeessa	hyperbole	ግነት
43	alxaa	appetite	አፒታይት

Declaration

I, the undersigned, declare that this dissertation is my own original work. This thesis has not been presented for any academic study in any other university, and all sources of material used for this work are clearly acknowledged.

Name: Desalegn Leshyibelu

Signature: -----

Date: June 2021

Place: Addis Ababa