

# The Similarity and Mutual Intelligibility between Amharic and Tigrigna Varieties

**Tekabe Legesse Feleke**

Verona University

Verona, Italy

hitakeleg@gmail.com

## Abstract

The present study has examined the similarity and the mutual intelligibility between Amharic and two Tigrigna varieties using three tools; namely Levenshtein distance, intelligibility test and questionnaires. The study has shown that both Tigrigna varieties have almost equal phonetic and lexical distances from Amharic. The study also indicated that Amharic speakers understand less than 50% of the two varieties. Furthermore, the study showed that Amharic speakers are more positive about the Ethiopian Tigrigna variety than the Eritrean variety. However, their attitude towards the two varieties does not have an impact on their intelligibility. The Amharic speakers' familiarity to the Tigrigna varieties seems largely dependent on the genealogical relation between Amharic and the two Tigrigna varieties.

Keywords: Language Similarity, Language Distance, Mutual Intelligibility, Attitude, Language Contact

## 1 Introduction

### 1.1 Language in Ethiopia

More than 85 languages are spoken in Ethiopia (Demeke, 2001; Hetzron, 1972; Hetzron, 1977; Hudson, 2013). The languages are classified under four language families: Semitic, Cushitic, Omotic and Nilo-Saharan (Bender and Cooper, 1976; Demeke, 2001; Hornberger, 2002; Hudson, 2013). In each family, there are many related language varieties so that the speakers of one variety can sometimes communicate with the speakers of another variety in the same language family

without major difficulties (Demeke, 2001; Gutt, 1980). However, the similarity among the languages is often obscured by the attitude of the speakers since language is considered as a symbol of identity (Lanza and Woldemariam, 2008; Smith, 2008). Hence, there are cases where varieties of the same languages are considered as different languages (Hetzron, 1972; Hetzron, 1977; Hudson, 2013; Smith, 2008). Therefore, due to politics, sensitivity to ethnicity and the lack of commitment from the scholars, the exact number of languages in Ethiopia is not known (Bender and Cooper, 1976; Demeke, 2001; Leslau, 1969). Furthermore, except some studies for example, Gutt (1980) and Ahland (2003) cited in Hudson (2013) on the Gurage varieties, and Bender and Cooper (1971) on mutual intelligibility of Sidamo dialects, the degree of mutual intelligibility among various varieties and the attitude of the speakers towards each others' varieties has not been thoroughly investigated. Hence, the present study examined the distance and the mutual intelligibility between Amharic and two Tigrigna varieties together with the effect of the attitude on the mutual intelligibility.

Amharic and Tigrigna are members of the Ethiosemitic language family, a branch of proto-Semitic family (Bender and Cooper, 1976; Demeke, 2001; Hetzron, 1972; Hetzron, 1977). According to Demeke (2001), Hetzron (1972), Hetzron (1977) and Bender and Cooper (1971), Ethiosemitic languages are divided into North and South Ethiosemitic. While the Tigrigna varieties are North Ethiosemitic languages, Amharic is one of the South Ethiosemitic languages. Nowadays, Amharic is spoken only in Ethiopia, but Tigrigna is spoken both in Ethiopia and in Eritrea. Due to the genealogical and typological relationship between Amharic and Tigrigna (Demeke, 2001; Hetzron, 1972; Hetzron, 1977), Amharic speak-

ers are supposed to understand the Tigrigna varieties to a certain degree. Since Amharic has been the national language of Ethiopia, it is a widely used language compared to Tigrigna (Getachew and Derib, 2008; Iyob, 2000; Lanza and Woldemariam, 2008; Smith, 2008). The use of Amharic as a national language helped many speakers of Ethiopian Tigrigna to learn Amharic as a second language (Smith, 2008). Moreover, Amharic has also been given as a subject for Ethiopian Tigrigna speakers, starting from elementary school. Some speakers of Eritrean Tigrigna variety used to speak Amharic before secession. However, after the independence, using Amharic in schools and in different offices was banned (Hailemariam and Walters, 1999; Rena, 2005). The relationship between the peoples of the two countries was also strained especially after Ethio-Eritrean war from 1988 to 2000. Hence, due to the border conflict, Eritrean Tigrigna speakers do not also have an access to Tigrigna speakers in Ethiopia and to the Amharic speakers.

Studies on the language attitude of the speakers of Amharic and the Tigrigna varieties are at scarce. However, language, ethnicity and politics are very interrelated in Ethiopia (Bulcha, 1997; May, 2011; Smith, 2008). The link has been accelerated by the ethnic-based federal system in Ethiopia (Lanza and Woldemariam, 2008; Young, 1996; Vaughan and Tronvoll, 2003). The atmosphere of politics in Eritrea and Ethiopia could also affect the attitude of the people in both countries. There has been an anti-Ethiopia sentiment in Eritrea since 1993 (Abbink, 2003; Assefa, 1996; Iyob, 2000). This hostile situation could have an effect on the attitude of the speakers of Amharic and the speakers of the Ethiopian Tigrigna.

The study of the similarity between Amharic and the Tigrigna varieties and the attitude of the speakers of one language towards another has a paramount significance in two ways. From practical point of view, there has been an attempt to standardize Tigrigna and use it widely in media and in schools. The study positively contributes to this effort. From theoretical perspective, there have been a number of attempts towards improving the enduring limitations of methods of dialectology. One of the positive contributions has been complementing the traditional lexicostatistics methods by the mutual intelligibility and perceptive distance measures. Very promising results

have been reported by the studies conducted on the Scandinavian languages and the Chinese dialects in this regard (see (Gooskens and Heeringa, 2004; Gooskens, 2013; Gooskens, 2007; Tang and Heuven, 2007; Tang and Heuven, 2009; Tang and Heuven, 2015)). The present study is an addition to these contributions.

## **1.2 Measuring Language Distance and Mutual Intelligibility**

The study of the distance among related languages has been a concern of many scholars for decades (Sokal, 1988). Several previous studies employed phonetic distance to measure the relative distance between various languages (Bakker, 2009; Cohn and Fienberg, 2003; Kessler, 1995). However, the emergence of the Levenshtein algorithm has enhanced the objective structural comparisons by introducing a computer-based distance computation (Heeringa, 2004; Gooskens and Heeringa, 2004). This has probably contributed a lot in terms of attracting many scholars towards the study of language variation (Gooskens, 2013). Recently, several studies have been conducted on European languages and on Chinese dialects, for example, (Gooskens and Heeringa, 2004; Heeringa, 2004; Tang and Heuven, 2007; Tang and Heuven, 2009; Tang and Heuven, 2015) by employing the Levenshtein algorithm together with the mutual intelligibility and perceptive distance measures. For instance, Gooskens (2007) compared data from Scandinavian languages (Danish, Swedish and Norwegian) with that of West Germanic languages (Dutch, Frisian and Afrikaans) and reported that mutual intelligibility can be predicted based on phonetic and lexical distances. Similarly, Bezooijen and Gooskens (2007) investigated the intelligibility of written Afrikaans and Frisian texts for Dutch speakers and reported the association between the Levenshtien distance and mutual intelligibility. Heeringa (2004) also employed the Levenshtein distance for the comparison of Dutch and Norwegian varieties.

The subjective measures often include perceptive distance and functional tests (Gooskens and Heeringa, 2004; Tang and Heuven, 2007; Tang and Heuven, 2009). According to Gooskens (2013), functional intelligibility between related languages can be measured by employing content questions, translation, recorded text testing, observations and performance tasks. Tang and Heuven (2009) employed word intelligibility test

and word recognition in a sentence to examine the mutual intelligibility among the Chinese dialects. According to Gooskens (2013) and Tang and Heuven (2009), opinion test can be designed without speech. For example, speakers of a certain variety can be requested to give their judgment on the speakers of other varieties who live in certain geographical areas (estimated linguistic distance). Bezooijen and Gooskens (2007) used cloze test to measure the functional ineligibility of written Afrikaans and Frisian for the native speakers of Dutch. Swarte and Gooskens (2014) employed a word translation to measure the importance of German for Dutch speaker to understand the Danish language. Hence, in the present study, the Levenshtein distance and the lexical distance were combined with intelligibility measure to determine the distance and the degree of intelligibility between Amharic and the Tigrigna varieties. Only the intelligibility of the Tigrigna varieties for the native speakers of Amharic was examined; the intelligibility measure was one directional primarily since measuring the degree of intelligibility of Amharic to the Tigrigna speakers was difficult as many Tigrigna speakers have an access to Amharic.

To measure the phonetic and the lexical distances, a written short story 'The Baboon Chief' was transcribed using IPA. Amharic and both Tigrigna varieties use Ethiopic writing system which means that there is a correspondance between the phonemes and graphemes; the difference is only on a few supra-segmental features which may not be captured in the written form. After the transcription, cognates in the stories were identified and aligned. Then, the distance between the cognates of each language was computed using Levenshtein distance. Lexical distance was determined by dividing non-cognate words to the total number of words in each text. Word translation was employed to measure the mutual intelligibility between the languages. Word translation was used since it was suitable for on line administration. Due to the complex socio-political situation in Ethiopia, the attitude of Amharic speakers towards the two language varieties and the contact between Amharic speakers and the speakers of the two Tigrigna varieties were also examined. Questionnaire was employed since it is suitable for on line administration (Agheyisi and Fishman, 1970). Bezooijen

and Gooskens (2007) also used questionnaires to examine the language contact and language background of their participants.

## 2 Research objectives

The study was conducted to address, among others, the following four specific objectives. 1) To determine the distance between written Tigrigna varieties and Amharic. 2) To determine the attitude of the native Amharic speakers towards the Tigrigna varieties. 3) To identify which Tigrigna variety is more intelligible for the native speakers of Amharic. 4) To indicate the relationship between the attitude of the speakers and the degree of mutual intelligibility.

## 3 Method

### 3.1 Participants

The participants were 18 native Amharic speakers who were attending MA program at Groningen, Rotterdam and Wageningen universities. Four of them were females and the remaining 14 were males. The average age of the participants was 27 year. Students who lived outside Ethiopia for more than two years were not included in the study since the attrition of Amharic could affect their responses and their performances on the mutual intelligibility test. Those whose parents are from Tigray Regional State-where Tigrigna is spoken or from Eritrea were also not included in the study; all of them were working in different colleges in Ethiopia before joining the three universities. The attitude and contact questionnaires were sent to each participant via email.

### 3.2 Materials and Tests

To measure the phonetic distance, the lexical distance and the intelligibility of the two Tigrigna varieties for the native speakers of Amharic, a fable 'The Baboon Chief' was translated from Oromo to Amharic by the researcher who is a balanced bilingual. The translation was checked by two independent bilingual experts (one is Oromo language instructor at Haromaya University and the second one Amharic instructor at Mekelle University). The selection of the fable from Oromo was to minimize the priming effect that could happen if it were taken directly from Amharic, see Tang and Heuven (2009) for the priming effect. The Amharic version was translated to the two Tigrigna varieties. The translators were native

speakers of the two varieties who were MA students at University of Groningen. The translated texts were checked by Tigrigna experts.

### 3.2.1 The Phonetic and Lexical Distance

The distances between Amharic and the two Tigrigna varieties were examined at two linguistic levels: phonetic and lexical. For the phonetic distance, the Levenshtein distance was employed. To apply the Levenshtein distance, the words in the translated written texts were transcribed using IPA symbols. To compute the phonetic distance, cognates both in Amharic and in the two Tigrigna texts were aligned. The distance between the corresponding cognates were determined based on a number of symbols which are inserted, deleted or substituted. The method of costs assignment was adopted from Gooskens (2007) with just a minor modification. The cost assignment is as follows: insertions and deletions 1 point, identical symbols 0 points, and substitutions of a vowel by a vowel or of a consonant by a consonant 0.5 point, substitutions of a vowel by a consonant or a consonant by a vowel 1 point. Below is an example of cost assignment which shows the distance between the cognates of Ethiopian Tigrigna and Amharic. In this example, the total cost (0.5 + 0.5) is one (1). The phonetic distance is the ratio of the total cost to the number of alignment (in this case 6). Thus, the phonetic distance is one divided by six (1/6) which is 0.167. In terms of percent, the distance between Tigrigna and Amharic cognates in this particular example is 16.7%.

k	u	l	l	o	m	Tigrigna
h	u	l	l	u	m	Amharic
.5	0	0	0	.5	0	Cost

The lexical distance between the two Tigrigna varieties and Amharic was determined based on the percentage of non-cognates in the total lexical items; the number of non-cognate words was divided to the total number of lexical items, based on Gooskens (2007). The cognates were identified based on two parameters which were suggested in Gooskens (2007): words in corresponding texts with common roots and cognate synonyms-words which are very similar in written form, but have slight meaning differences (e.g. *hajal* 'powerful' and *hajl* 'power'). Whether the pairs of words are cognates or not was determined by two Amharic and Ethiopian Tigrigna bilinguals and another two

Amharic-Eritrean Tigrigna bilinguals.

### 3.2.2 Language Attitude and Contact

To examine the attitude of the Amharic speakers towards the two Tigrigna varieties, questionnaires were adopted from Bezooijen and Gooskens (2007). The questionnaires contained items which focus on the two Tigrigna varieties, on the speakers of the varieties and on areas where the two Tigrigna varieties are spoken. For each area of interest, three items and the total of eighteen items were constructed. The participants provided their responses on the items that contain five point scales. For example, they rate whether Tigrigna is an interesting language or not on the scale: 1 (interesting) to 5 (extremely boring).

Similarly, questionnaires were employed for the assessment of the participants' contact with the two Tigrigna varieties. The questionnaires included items related to the participants' frequency of contact with the speakers of the two varieties, media, movies, and newsletters of the two varieties. The participants rated the degree of contact by using five rating scales (very often, often, occasionally, very rarely and not at all). Ten questions were provided for each variety, and the items designed to measure each variety were evenly distributed.

### 3.2.3 The Mutual Intelligibility

Word translation was used for the mutual intelligibility measure due to its ease of administration. In the translation task, Tigrigna words in the translated fable were listed based on alphabetical order, and 100 words from each Tigrigna variety, a total of 200 words were selected. Since there were 136 Eritrean and 130 Ethiopian words in the translated texts, 36 words from Eritrean Tigrigna texts and 30 words from Ethiopian Tigrigna texts were randomly left out, and the remaining 100 words in each text were used for the test. Since translating 200 words could be a tiresome task for the participants, the 200 words in the two Tigrigna varieties were divided across the participants. Hence, among 18 participants who took part in the test, nine participants translated the first 50 words in the lists of each of the varieties, and the remaining nine participants translated the last 50 words in the lists of each variety. Using this procedure, each participant translated 100 words (50 from each variety) to Amharic. Since translating the list of words of one variety and shifting to the list of

words of another variety could lead to priming (see Tang and Heuven (2009)), words from the two varieties were mixed, but were written in a slightly different font so that the researcher could identify to which variety each word belongs.

Then, the mixed words were evenly distributed in such a way that each translator received different word order. To achieve this, the mixed 100 words were initially randomly ordered and numbered. Then, different word orders were created using base ten as a point of classification. In this manner, the first order begins with No.1 and ends with No.100 (the default order). The second order begins with No.10 followed by from 11-100 and then from 1-9. The third order begins with No.20 followed by from 21-100 then from 1-19 and so on. In this manner nine different order for each group, and the total of 18 list of orders were created. The respondents were instructed to translate each word within 30 seconds. However, it is important to recognize that the participants could take less or more than the allotted time since the task was administered on line. The intelligibility measure is the number of words which was translated correctly. One (1) point was given for fully correct answer, and 0.5 point was given for correct answers but with tense, aspect, number and other morphological/grammatical errors. The appropriateness of the translation was checked by the researcher and by the native speakers of the two varieties.

## 4 Results

### 4.1 The Phonetic Distance

The two Tigrigna varieties have about 30% phonetic differences with Amharic. In other words, the two varieties have equal phonetic distance from Amharic; Ethiopian Tigrigna (M = 31%) and Eritrean Tigrigna (M = 28.5%); independent t-test,  $t = .023$ ,  $p = .56$ . Among 136 total words, there were 51 Eritrean Tigrigna cognate words, and 85 non-cognates words. Hence, the lexical distance between Amharic and Eritrean Tigrigna variety is 62.5% (85/136). This means that the lexical similarity between Amharic and the Eritrean Tigrigna variety is 37.5%. The Ethiopian Tigrigna text contains 130 words. Among these, 59 (43.5 %) were cognates, and 71 (56.3) were non-cognates. This shows that the lexical distance between Amharic and the Ethiopian Tigrigna variety is 45.4% (71/130). The results indicate that

Amharic is more closer to Ethiopian Tigrigna variety than to the Eritrean Tigrigna Variety.

### 4.2 Language Attitude and Language Contact

The Amharic speakers are more positive about Ethiopian Tigrigna (M = 3.5) than the Eritrean Tigrigna (M = 3), paired sample t-test,  $t = -2.754$ ,  $p = .01$ . The attitude of the Amharic speakers was also examined specifically in terms of the three areas of interest: attitude towards the language, attitude toward the people, and attitude towards the country. As Table 1 shows, Amharic speakers are more negative about Eritrea. The difference is significant in all cases except in their attitude towards the people (paired t-test  $t = .849$ ,  $p = 0.42$ ). With regard to the language contact, Amharic speakers have stronger contact to Ethiopian Tigrigna than to the Eritrean Tigrigna; paired sample t-test:  $t = -7.923$ ,  $p = .00$ . Nothing is surprising about this finding since the speakers of Amharic do not have a direct contact with the Eritrean Tigrigna speakers due to the border conflict between the two countries. Though the contact between the speakers of Amharic and the speakers of Ethiopian Tigrigna is higher than the contact between the speakers of Amharic and that of the speakers of Eritrean Tigrigna, it does not seem that Amharic speakers have a frequent contact with Ethiopian Tigrigna speakers as the frequency of contact is very low (2.9 on 1-5 scale).

Focus	ERT	ETT	t	Sig
Lang	3.3	3.9	3.1	.01
Peop	3.6	3.7	.85	.40
Coun	1.9	2.8	2.8	.02
Mean	3.0	3.5	-2.8	.01
Contact	1.8	2.9	-7.9	.00

The attitude of Amharic speakers towards the Tigrigna varieties measured on (1-5) Linker scale. 'ERT' refers to Eritrea Tigrigna, ETT refers to Ethiopian Tigrigna, 'Lang' is language, and 'Coun' refers to country.

### 4.3 Mutual Intelligibility

The mutual intelligibility test results indicate that Amharic speakers have equal performances on both languages; (M = 29.78%) on Ethiopian Tigrigna and (M = 26.11%) on Eritrean Tigrigna, paired sample t-test,  $p = .15$ . Besides, the atti-

tude and contact results do not correlate with intelligibility results,  $r = -.267$  and  $r = 0.181$  respectively. Likewise, there is no correlation between Amharic speakers' contact with the Eritrean Tigrigna speakers and their performance on the Eritrean Tigrigna mutual intelligibility test.

## 5 Discussion

Results obtained from the phonetic and the lexical distance measures show that the two Tigrigna varieties have almost equal distance from Amharic. The lexical distance between Amharic and the Ethiopian Tigrigna is also similar with the one between Amharic and the Eritrean Tigrigna. The speakers of Amharic are more negative about Eritrean Tigrigna variety than the Ethiopian Tigrigna variety. The negative attitude towards Eritrea is not astonishing since there was political and ethnic hostility between the two countries which might have affected the Amharic speakers attitude towards Eritrea and Eritrean Tigrigna (Hailemariam and Walters, 1999; Rena, 2005).

Though the attitude of the Amharic speakers is more positive towards Ethiopian Tigrigna, the magnitude of the attitude is not high. This can be due to political reasons since there has been a fierce power struggle between the Amhara and the Tigray ethnic groups (Young, 1998). Amharic speaker have a stronger contact with the Ethiopian Tigrigna speakers than with the Eritrean Tigrigna speakers. However, in both cases, the frequency of contact is low. As presented earlier, contacting the Eritrean people is almost impossible for the Amharic speakers as the communication between the two countries has been blocked due to the border conflict. The contact between the Amharic speakers and the Ethiopian Tigrigna speakers is also small. This could be due to economic, language and social situation in the country. Tigray region is found in the northern tip of the country, very distant place from the capital. Usually, people move from Tigray region to the central part of the country where Amharic is used to seek job, education, recreation and other purposes. There is a less possibility for Amharic speakers to move to Tigray region.

The results obtained from the intelligibility test show that both Tigrigna varieties are almost equally difficult to the native Amharic speakers. This can have two possible interpretations. In one hand it shows that the two Tigrigna varieties

have almost equal distance from Amharic. On the other hand, it indicates that native Amharic speakers cannot communicate with the speakers of both Tigrigna varieties using Tigrigna as a medium of communication since the Amharic speakers scored below the average on the mutual intelligibility tests. According to Gutt (1980), two languages are considered as intelligible if the speakers of one variety understand more than 80% of another variety. This means that the two Tigrigna varieties are not intelligible for the native Amharic speakers. Besides, Amharic speakers' intelligibility scores on both language varieties are not affected by both language contact and attitude. This finding is consistent with that of Bezooijen and Gooskens (2007) and Gooskens and Heeringa (2004) that there may not be a correlation between language attitude and language intelligibility. The absence of correlation between language contact and language mutual intelligibility shows that the distance and the magnitude of intelligibility which were reported in the present study are due to the genealogical relationship between Amharic and the two Tigrigna varieties.

In general, this study indicates that both the Ethiopian and the Eritrean Tigrigna varieties have almost a comparable phonetic and lexical distance from Amharic. Native Amharic speakers understand less than half of the two varieties which hints that the two Tigrigna varieties are not intelligible for the Amharic speakers. Furthermore, the speakers of Amharic are more positive about the Ethiopian Tigrigna variety than the Eritrean variety. Nevertheless, their attitude does not have an impact on their intelligibility of the two varieties. Moreover, the study has shown that Amharic speakers have more frequent contact with the Ethiopian Tigrigna speakers than with the Ethiopian Tigrigna speakers. However, their familiarity to the two Tigrigna varieties has nothing to do with the contact between the speakers of the two languages.

The present study is perhaps the first attempt towards establishing the mutual intelligibility and measuring the relative distance between Amharic and the two Tigrigna varieties. Future studies ought to consider a large scale research which includes all the Amharic and the Tigrigna dialects.

## References

- Abink. 2003. Badme and the ethio-eritrean border: the challenge of demarcation in the post-war period. *Africa: Rivista trimestrale di studi e documentazione dell'Istituto italiano per l'Africa e l'Oriente*, 58(2):219–231.
- Agheysi and Fishman. 1970. Language attitude studies: A brief survey of methodological approaches. *Anthropological linguistics*.
- Ahland. 2003. Interlectal intelligibility between gurage speech varieties. In *In North Conference on Afroasiatic Linguistics*, April.
- Assefa. 1996. Ethnic conflict in the horn of africa: myth and reality. *Ethnicity and power in the contemporary world*.
- Bakker. 2009. Adding typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology*, 13(1):169–181.
- Bender and Cooper. 1971. Mutual intelligibility within sidamo. *Lingua*, 27(32-52).
- Bender and Cooper. 1976. Language in ethiopia: Implications of a survey for sociolinguistic theory and method. *Pub date notes*, pages 75–191.
- Bezooijen and Gooskens. 2007. Interlingual text comprehension: linguistic and extralinguistic determinants. *Hamburger Studies in Multilingualism*, (249-264).
- Bulcha. 1997. The politics of linguistic homogenization in ethiopia and the conflict over the status of afaan oromo. *African affairs*, 96(384):325–352.
- Ravikumar Cohn and Fienberg. 2003. A comparison of string distance metrics for name-matching task. . In *IWeb*.
- Demeke. 2001. The ethio-semitic languages (re-examining the classification). *Journal of Ethiopian Studies*, pages 57–93.
- Getachew and Derib. 2008. Language policy in ethiopia: History and current trends. *Ethiopian journal of education and sciences*, 2(1).
- Gooskens and Heeringa. 2004. Perceptive evaluation of levenshtein dialect distance measurements using norwegian dialect data. *Language variation and Change*, 16(03):189–207.
- Gooskens. 2007. The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and multicultural development*, 28(6):445–467.
- Gooskens. 2013. *Experimental Methods for Measuring Intelligibility of Closely Related Language Varieties*. Oxford University Press.
- Gutt. 1980. Intelligibility and interlingual comprehension among selected gurage speech varieties. *Journal of Ethiopian Studies*, 14:57–85.
- Kroon Hailemariam and Walters. 1999. Multilingualism and nation building: Language and education in eritrea. *Journal of Multilingual and Multicultural Development*, 20(6):457–493.
- Heeringa. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Ph.D. thesis, University of Groningen.
- Hetzron. 1972. *Ethiopian Semitic: studies in classification*. Manchester University Press.
- Hetzron. 1977. *The Gunn-Gurage Languages*, volume 2. Istituto orientale di Napoli.
- Hornberger. 2002. Multilingual language policies and the continua of biliteracy: An ecological approach. . *Language policy*, 1(1):27–51.
- Hudson. 2013. *Northeast African Semitic: Lexical Comparisons and Analysis*. Harrassowitz.
- Iyob. 2000. The ethiopianeritrean conflict: diasporic vs. hegemonic states in the horn of africa, 19912000. *The Journal of Modern African Studies*, 38(04):659–682.
- Kessler. 1995. Computational dialectology in irish gaelic. In *In Proceedings of the seventh conference on European chapter of the Association for Computational Linguistics*. Morgan Kaufmann Publishers Inc..
- Lanza and Woldemariam. 2008. Language policy and globalization in a regional capital of ethiopia. In *Linguistic landscape*. Expanding the scenery.
- Leslau. 1969. Toward a classification of the gurage dialects. *Journal of Semitic Studies*, 14(1):96–109.
- May. 2011. *Language and minority rights: Ethnicity, nationalism and the politics of language*. Routledge.
- Rena. 2005. Eritrean education-retrospect and prospect. *Journal of Humanities and Sciences*, 5(2):1–12.
- Smith. 2008. The politics of contemporary language policy in ethiopia. *Journal of Developing Societies*, 24(2):207–243.
- Sokal. 1988. Genetic, geographic, and linguistic distances in europe. In *Proceedings of the National Academy of Sciences*, volume 85 of 5, pages 1722–1726.
- Schppert Swarte and Gooskens. 2014. Does german help speakers of dutch to understand written and spoken danish words? the role of second language knowledge in decoding an unknown but related language. in press.

- Tang and Heuven. 2007. Mutual intelligibility and similarity of chinese dialects: Predicting judgments from objective measures. *Linguistics in the Netherlands*, 24(1):223–234.
- Tang and Heuven. 2009. Mutual intelligibility of chinese dialects experimentally tested. *Lingua*, 119(5):709–732.
- Tang and Heuven. 2015. Predicting mutual intelligibility in chinese dialects from subjective and objective linguistic similarity. *Interlinguistica*, 17:1019–1028.
- Vaughan and Tronvoll. 2003. *The culture of power in contemporary Ethiopian political life*. Sida, Stockholm.
- Young. 1996. Ethnicity and power in ethiopia. *Review of African Political Economy*, 23(70):531–542.
- Young. 1998. Regionalism and democracy in ethiopia. *Third World Quarterly*, pages 191–204.