

Title	The $\it n$ sound in Cushitic and Omotic languages
Author(s)	Wakasa, Motomichi
Citation	スワヒリ&アフリカ研究. 2023, 34, p. 26-38
Version Type	VoR
URL	https://doi.org/10.18910/91112
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

The *n* sound in Cushitic and Omotic languages

Wakasa, Motomichi*

0. Introduction

Wolaytta and Kambata are neighboring languages in the southwestern part of Ethiopia. Although both are said to be Afroasiatic languages (some disagree on this point), the two languages belong to different families—Wolaytta is an Omotic language and Kambata a Cushitic language—and they are not mutually intelligible because of their great differences in vocabulary and grammar.

According to Wakasa (2016), however, they share an interesting phonotactic tendency: the markedness of the n sound against the m sound in word-initial position (see section 1.).

This paper surveys other Cushitic and Omotic languages on this matter using published materials.¹

1. The markedness of the n sound in Wolaytta and Kambata

1.1 Wolaytta

Wolaytta has a voiced dental nasal, n, as a phoneme. This sound appears relatively frequently in texts, partly because it is found in some verbal endings, postpositions, and basic words such as personal pronouns. Thus, the n sound is not rare in this language.

However, words that begin with the *n* sound are relatively rare, while those with an initial voiced bilabial nasal, *m*, are not. For example, in Alemaayehu and Tereezzaa's (1991 EC) Wolaytta-Amharic dictionary, headwords with initial *n* start at the end of page 236, and end at the end of page 240. In other words, they occupy only 5 pages (actually 4 pages and a few lines) out of 421. In this dictionary, of course, many derived forms appear as headwords. If such derivatives are removed, by my count, only 48 morphemes with initial *n* can be recognized. Furthermore, these morphemes include loans from Amharic, onomatopoeia (or ideophones), and morphemes which do not seem to be used in ordinary life (at least, I have not encountered them outside of this dictionary).

This is not the case for the m sound. In the same dictionary, the headwords with initial m start at the middle of page 207, and end at the end of page 236. In other words, they occupy 30 pages (net 29 pages or so) out of 421. Thus, the initial voiced bilabial nasal is about six times as numerous as the initial voiced dental nasal.

^{*} Part-time lecturer at Atomi University / Hakuoh University / Tokyo University of Foreign Studies / Meisei University

When referring to the names of the languages in this paper, I follow Hayward (2000).

The same tendency can be ascertained in other materials. In Lemma (1992 EC), the headwords with initial m start at the end of page 91, and end at the end of page 108. Those with initial n start at the beginning of page 109, and end at the end of page 112. This dictionary has 199 pages in all. In Wakasa (2000), the headwords with initial m start at the middle of page 93 and end at the beginning of page 97, and those with initial n occupy the rest of that page. Again, the m sound is about four times as popular as the n sound.

As noted above, neither m nor n is a rare sound in Wolaytta. Furthermore, according to Jakobson ([1941=] 1969: 61–62), these two sounds are constituents of the minimal consonantal system of the languages of the world. In other words, m and n are expected to be found in every language unless some special conditions prevent them from occurring, and are expected to be acquired in the earlier stages of language acquisition by infants. Thus, it is surprising to find a tendency, if not a restriction, regarding the environment in which one of these "basic and simple" sounds occurs in Wolaytta.

1.2 Kambata (Kambaata)

Kambata also has a voiced dental nasal, n, as a phoneme. In this language, words that begin with the n sound are even rarer than in Wolaytta. I was able to find only 9 such morphemes, including loans, during my fieldwork on Kambata, while I was able to find about 140 morphemes with initial m.

Specifically, the morphemes I found beginning with n are the following: $naaqar-\dot{u}ta$ '(a kind of spice)', $naar-\dot{u}$ 'to pile', $nagaariit-\dot{a}$ 'drum', $naqqas-\dot{u}$ 'to testify', $naxib-\dot{a}$ 'point', $n\acute{e}\acute{e}si$ 'we', $nig-\dot{a}$ 'tendon', $nubaab-\dot{u}$ 'old man', $nugguss-\dot{u}$ 'to become pregnant', $n\acute{u}ri$ 'last year'. The postposition -n 'in, at, by' and the conjunctive $-n\acute{u}$ 'and' might be added here, though these never occur at the beginning of a phrase.

According to Kazuhiro Kawachi [河内一博] (personal communication), the *n* sound is similarly marked in Sidamo (Sidaama), which is a Cushitic language genetically close to Kambata and one of the neighboring languages of Wolaytta.

What is the situation in other Cushitic and Omotic languages? The following sections try to shed some light on this question.

2. Research method

2.1 Classification of Cushitic and Omotic languages

There have been controversies over genetic classification and historic development of Cushitic and Omotic languages. I will not go into these matters here. Instead, I will follow Hayward's (2000) classification, since this work seems to be the most standard and recent concise introduction to the Afroasiatic languages. His classification can be summarized as follows (Table 1):

Table 1 Hayward's (2000) classification of Cushitic and Omotic languages

Cushitic

1. Northern Cushitic: Bedawi/Beja

2. Central Cushitic (Agaw): Bilin, Kemant, Kwara, Xamtanga, Awngi

3. Highland East Cushitic: Burji, Sidamo, Kambata, Hadiyya

4. Lowland East Cushitic: (1) (northern): Saho, Afar

(2) (Oromoid): Oromo, Konso

(3) (Omo-Tana, Eastern): Rendille, Boni, Somali

(Omo-Tana, Western): Daasenech, Arbore, Elmolo

(Omo-Tana, Isolated): Bayso

5. Dullay: Tsamay, Gawwada

6. Southern Cushitic: Iraqw, Gorowa, Burunge, Mbugu/Ma'a, Asax, Kw'adza, Dahalo.

Omotic

South Omotic: Aari, Hamer-Banna, Karo, Dime **North Omotic:** (1) **Dizioid:** Dizi, Nayi, Sheko

(2) Gonga-Gimojan: Gonga: Kaficho, Shakacho, Boro, Anfillo

Gimojan: Yemsa

Gimira-Ometo: Gimira: Bench

Ometo: Wolaytta, Gamo, Gofa,

Basketto, Male, Chara

Mao: Eastern: Bambassi Western: Hozo, Seze

2.2 Research method adopted in this study

For each of the languages selected for examination from Table 1, I performed the following procedure, using published materials.

(A) I counted the number of phonemes and divided 100 by the result to calculate an "expected value." The expected value is an average that corresponds to the percentage of words that would begin with a given phoneme in the absence of phonotactic tendencies.

- (B) I calculated the actual percentages of words with initial m and n in the vocabulary, respectively.
- (C) I compared the result from (B) for initial *n* with the expected value calculated in (A).

(D) I compared the result from (B) for initial n with the result from (B) for initial m.

2.3 Problems with the method adopted in this study

The procedure described in section 2.2 has many problems. First, I used published materials. This means that I was not able to use firsthand data. Thus, I cannot evaluate the correctness of the materials I used.

The materials also differ in size. Some are relatively small—perhaps too small for this kind of quantitative research. They also differ in their aim. Some are dictionaries, while others are just word lists. Some are for synchronic description, while others are aimed at historical comparison. I am not sure whether I understand and appreciate their contents in all cases.

For some shorter materials I was able to count the number of items. For relatively bulky dictionaries, I only counted the number of pages. For Wolaytta and Kambata I was able to count morphemes since I know the languages, but for the other languages I was obliged to count headwords mechanically since I could not judge what the headwords represented—morphemes, derived stems, or inflected forms?

Counting the number of phonemes in a language is not easy because it involves interpretation, which can be sometimes subjective. For example, should we consider initial glottal stops to be phonemic or should we ignore them considering them not to be phonemes? (I counted in both ways.) Should we count a long vowel as a single separate phoneme or as a succession of two identical short vowel phonemes? (The latter is my choice in this work.)

Morphemes that never occur word-initially, such as suffixes and postpositions, should be excluded from the count in such studies as this. But I may have included them when counting, because of my insufficient understanding of the languages in question.

In short, the methodology used in this work is far from rigorous. However, I believe it does permit a valuable preliminary survey of general tendencies regarding the markedness of word-initial m and n.

3. Results

3.1 Sources and data

This section lists the languages surveyed here, the sources used, and the figures obtained by the methodology described in section 2.2. Figures marked (A) represent the expected value for (percentage of) words that would begin with each phoneme in the language's inventory in the absence of phonotactic tendencies, and figures marked (B) the actual percentages of words with initial m and n. The figures from steps (C) and (D) in section 2.2 are omitted since they can easily be calculated from the following data.

Bedawi/Beja: Reinisch (1895) (A) 4% (4.76%²)

(B) *m*: 6:61%, *n*: 5.37%

Bilin: Appleyard (2006)

(A) 2.78% (3.45%)

(B) *m*: 4.97%, *n*: 3.39%

Xamtanga: Appleyard (1987)

(A) 2.78% (3.13%)

(B) *m*: 3.51%, *n*: 2.11%

Awngi: Appleyard (2006)

(A) 2.86% (3.45%)

(B) *m*: 4.66%, *n*: 1.60%

Kambata: My own fieldnotes

(A) 3.70% (4.35%)

(B) *m*:10.84%, *n*: 0.68%

Afar: Parker and Hayward (1985)

(A) 4.55% (5.56%)

(B) *m*: 8.56%, *n*: 2.14%

Oromo: Tilahun (1989)

(A) 3.33% (3.85%)

(B) *m*: 6.59%, *n*: 1.48%

Konso: Kowaki (2008)

(A) 3.85% (4.76%)

(B) *m*: 6.31%, *n*: 1.52%

Somali: Puglielli (et al.) (1985)

(A) 3.87% (4.55%)

(B) *m*: 9.10%, *n*: 2.31%

Daasenech: Tosco (2001)

(A) 3.33% (3.85%)

(B) *m*: 5.71%, *n*: 2.62%

Bayso: Hebarland and Lamberti (1988)

(A) 3.45% (4.00%)

(B) m: 6.90%, n: 1.47%

² For all languages, the figures in parentheses are the expected values when initial vowels are regarded as preceded by the glottal stop.

Tsamay: Savà (2005)

(A) 2.94% (3.33%)

(B) m: 7.21%, n: 1.47%

Iraqw: Maghway (1995)

(A) 3.03% (3.45%)

(B) *m*: 4.55%, *n*: 4.53%

Dahalo: Tosco (1991)

(A) 1.85% (2.00%)

(B) *m*: 6.58%, *n*: 2.36%

Aari: Tsuge (1995)

(A) 3.57% (4.17%)

(B) *m*: 4.17%, *n*: 3.33%

Hamer: Takahashi (2009)

(A) 3.33% (4.00%)

(B) *m*: 3.02%, *n*: 3.36%

Dime: Tsuge (2008)

(A) $2.94\% (3.57\%)^3$

(B) *m*: 5.26%, *n*: 2.42%

Sheko: Hellenthal (2010)

(A) 2.94% (3.45%)

(B) *m*: 3.77%, *n*: 2.78%

Shakacho: Leslau (1959)

(A) 3.57% (4.76%)

(B) *m*: 8.93%, *n*: 2.52%

Boro: Lamberti (1993a)

(A) 3.03% (3.45%)

(B) *m*: 8.62%, *n*: 2.77%

Yemsa: Lamberti (1993b)

(A) 3.23% (3.70%)

(B) *m*: 9.62%, *n*: 2.76%

Bench: Wedekind (1990)

(A) 3.03% (3.33%)

(B) *m*: 5.18%, *n*: 3.78%

³ Number of phonemes taken from Fleming (2000).

Wolaytta⁴: Alemaayehu and Tereezza (1991 EC)

(A) 2.94% (3.33%) (B) *m*: 7.13%, *n*: 1.19%

Basketto: Inui (2009)

(A) 3.03% (3.45%) (B) *m*: 10:71%, *n*: 1.34%⁵

3.2 Interpretation of the data: Three types of languages in terms of markedness of the n sound

Using the figures from (C) and (D) described in section 2.2, we can classify Cushitic and Omotic languages into three types.

Type I (Kambata-Wolaytta type) consists of languages where words with initial n are relatively rare in the vocabulary and much rarer than those with initial m. Concretely speaking, it consists of languages where the percentage of words with initial n is less than half of the expected value and less than a third of the percentage with initial m. These languages are: **Kambata**, Afar, Oromo, Konso, Bayso, Tsamay, **Wolaytta**, *Basketto*.

Type II consists of languages where words with initial n are not notably rare in the vocabulary but still much rarer than those with initial m. Concretely speaking, it consists of languages where the percentage of words with initial n is more than half of the expected value but less than a third of the percentage with initial m. These languages are: Somali, Shakacho, Boro, Yemsa.

Type III (non-Kambata-Wolaytta type) consists of languages where words with initial n are neither notably rare in the vocabulary nor notably rarer than those with initial m. Concretely speaking, it consists of languages where the percentage of words with initial n is more than half of the expected value and more than a third of the percentage with initial m. These languages are: Bedawi/Beja, Bilin, Xamtanga, Awngi, Daasenech, Iraqw, Dahalo, Aari, Hamer, Dime, Sheko, Bench.

The criteria used here are, of course, arbitrary. As observed in notes 6 and 7, classification is difficult in some cases. The signification of each percentage differs from language to language because the number of phonemes differs from language to language. Some languages have η and/or η , which might be related to η historically. Nevertheless, even with all these problems taken into consideration, this grouping into three types is still justifiable as a guide to the phonotactic tendencies of the languages

_

⁴ See also the discussion in section 1.1.

⁵ If we use Inui (2002), however, the figures are 9.79% (*m*) and 2.37% (*n*).

⁶ It is evident from the figures that Kambata and Wolaytta, for which I happen to have firsthand data, are particularly typical members of this type (thus set in boldface type). Basketto seems to belong to this type, but the figures seen in Inui (2002) do not testify any great rarity of initial *n* (it is thus set in italic type).

Awngi is a borderline case. It does not meet the criteria for Type III if we use the expected value in the parentheses, rather than outside the parentheses, in section 3.1 above. Furthermore, the percentage for initial n is almost a third of the percentage for initial m, barely meeting that criterion. Thus, the language is set in italic type.

under discussion.

3.3 From a genetic viewpoint

Table 2 shows the results from section 3.2 combined with the tree in Table 1, which reflects a standard opinion on the genetic relationship of Cushitic and Omotic languages. Type I languages are colored green, Type II yellow, and Type III red. Languages not marked with any color are those not included in this survey.

Table 2

Cushitic

1. Northern Cushitic: Bedawi/Beja

2. Central Cushitic (Agaw): Bilin, Kemant, Kwara, Xamtanga, Awngi

3. Highland East Cushitic: Burji, Sidamo, Kambata, Hadiyya

4. Lowland East Cushitic: (1) (northern): Saho, Afar

(2) (Oromoid): Oromo, Konso

(3) (Omo-Tana, Eastern): Rendille, Boni, Somali

(Omo-Tana, Western): Daasenech, Arbore, Elmolo

(Omo-Tana, Isolated): Bayso

5. Dullay: Tsamay, Gawwada

6. Southern Cushitic: Iraqw, Gorowa, Burunge, Mbugu/Ma'a, Asax, Kw'adza, Dahalo.

Omotic

South Omotic: Aari, Hamer-Banna, Karo, Dime

North Omotic: (1) Dizioid: Dizi, Nayi, Sheko

(2) Gonga-Gimojan: Gonga: Kaficho, Shakacho, Boro, Anfillo

Gimojan: Yemsa

Gimira-Ometo: Gimira: Bench

Ometo: Wolaytta, Gamo, Gofa,

Basketto, Male, Chara

Mao: Eastern: Bambassi
Western: Hozo, Seze

As Table 2 shows, it is difficult to draw any decisive conclusion from the viewpoint of genetic affiliation. For example, a Type I language may be Cushitic or Omotic, while an Omotic language may be Type I, II, or III.

3.4 From a geographical viewpoint

Map 1 shows the results from section 3.2 plotted on a map indicating where each Cushitic or Omotic language is mainly spoken.⁸ The languages that the numbers stand for are as follows: 20 Bedawi/Beja, 21 Bilin, 22 Xamtanga, 24 Awngi, 26 Afar, 28 Oromo, 29 Konso, 30 Somali, 33 Bayso, 34 Daasenech, 38 Kambata, 45 Dullay (Tsamay is a southern variety of this linguistic chain), 47 Iraqw, 48 Dahalo, 55 Wolaytta, 64 Basketto, 71 Bench, 73 Yemsa, 75 Mocha (=Shakacho), 76 Boro, 79 Sheko, 85 Hamer, 88 Aari, and 89 Dime.

Here again, Type I languages are colored green, Type II yellow, and Type III red. Languages not marked with any color are those not included in this survey. Wolaytta and Kambata have darker colors and Basketto and Awngi lighter ones for the reasons explained in notes 6 and 7, corresponding to use of boldface and italic type.⁹

Very roughly speaking, Map 1 shows that Kambata and Wolaytta are surrounded by Type I languages, Type I languages are neighbored by Type II languages, and Type III languages are situated in places that are distant from the Kambata-Wolaytta region.

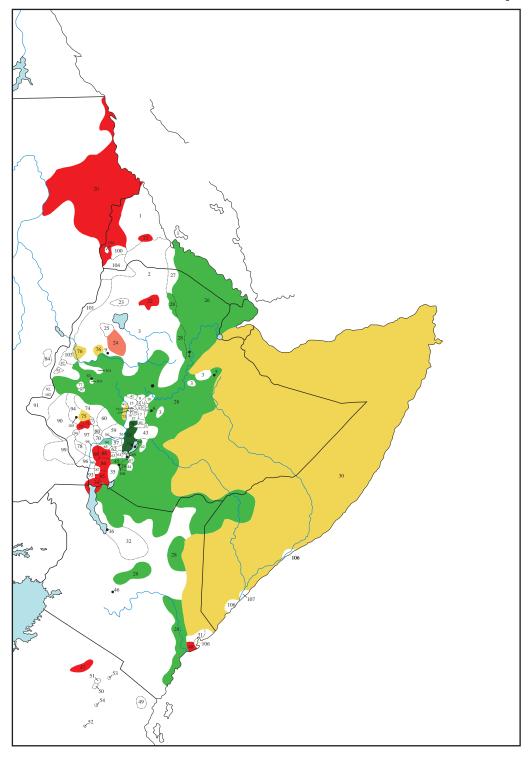
Geographical relationships thus explain the distribution of the markedness of the *n* sound better than genetic ones. The feature is found in Kambata, a Cushitic language, but is not shared by Northern, Central, or Southern Cushitic languages, which are spoken in places that are distant from the Kambata region. However, the same feature is shared by Wolaytta, a neighboring North Omotic language, and again, the feature is not shared by Sheko and Bench, North Omotic languages spoken in places that are distant from the Wolaytta region.

_

⁸ This map was made based on Map 77 in Moseley and Asher (1994) with invaluable contributions by Hiroki Ishikawa [石川博樹], associate professor of the Research Institute for Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, to whom I am deeply grateful.

⁹ Probably, Sidamo (number 43) can also be colored green. See section 1.2.

Map1



4. Final remarks

This study revealed the distribution of a phonological tendency, i.e., the markedness of the n sound in Cushitic and Omotic languages. It was not able to explain the distribution of this tendency, which is beyond the study's scope.

However, future comparative and/or historical studies on these languages will have to take these results under consideration somehow. It is especially interesting that South Omotic languages, which are argued by some to belong to the same family as Wolaytta and spoken relatively near the Wolaytta zone, are not members of Type I, i.e., the Kambata-Wolaytta type.

According to Shuichiro Nakao [仲尾周一郎] (personal communication), some Nilotic languages, such as Bari and Nandi, have many more initial-*m* words than initial-*n* ones. Thus, we may have to explain the tendency in terms of language contact.

Acknowledgement

My fieldtrips aimed at collecting the basic vocabulary of Kambata were supported by JSPS KAKENHI Grant Number 19720093 and 20320059.

References

("EC" stands for the Ethiopian Calendar.)

- Agostini, Francesco, Annarita Puglielli and Ciise Moxamed Siyaad (eds.). 1985. *Dizionario Somalo Italiano*. Rome: Gangemi Editore.
- Alemaayehu Doogamo and Tereezzaa Hayile Messqqalo. 1991 EC. *Wolayttatto qaalatu Amaaratto birshshettaa* [Wolaytta-Amharic dictionary]. Addis Ababa, Tophphiyaa Doonatu Xinaatiyaanne Pilggettaa Ooso Keettaa, Addisaaba Yuniversttiyaa.
- Appleyard, David L. 1987. "A grammatical sketch of Khamtanga I, II." *Bulletin of the School of Oriental and African Studies*, 50, 241–266, 470–507.
- Appleyard, David L. 2006. A comparative dictionary of the Agaw languages. Köln, Rüdiger Köppe Verlag.
- Fleming, Harold. 2000. "A grammatical sketch of Dime (Dim-Af)." Richard J. Hayward (ed.). *Omotic language studies*, 494–583. London, School of Oriental and African Studies, University of London.
- Hayward, Richard J. 2000. "Afroasiatic." Bernd Heine and Derek Nurse (eds.). *African languages: An introduction*, 74–98. Cambridge, New York, Melbourne, Madrid, Cambridge University Press.

- Haberland, Eike and Lamberti M. 1988. *Ibaaddo ka-Ba'iso*. *Culture and language of the Ba'iso*. Heidelberg, Carl Winter Universitätsverlag.
- Hellenthal, Anne-Christie. 2010. A grammar of Sheko. Utrecht, LOT.
- Inui, Hideyuki [乾秀行]. 2002. "Basuketogo no goi [バスケト語の語彙]" [A vocabulary of Basketto]. *Ippan gengogaku ronsō* [一般言語学論叢], 4 and 5, 10–32.
- Inui, Hideyuki [乾秀行]. 2009. "Basuketogo dōshi 500 [バスケト語動詞 500]" [500 Basketo verbs]. Hideyuki Inui (ed.). *Cushitic-Omotic studies 2008*, 11–93. Yamaguchi.
- Kowaki, Mitsuo [小脇光男]. 2008. "Konsogo no kisogoi mojika e no kokoromi [コンソ語の基礎語彙—文字化への試み—]" [A basic vocabulary of the Konso language: With a tentative Latin-based orthography]. Hideyuki Inui (ed.). *Cushitic-Omotic studies 2007*, 93–139. Yamaguchi.
- Lamberti, Marcello. 1993a. *Die Shinassha-Sprache. Materialien zum Boro*. Heidelberg, Universitätsverlag C. Winter.
- Lamberti, Marcello. 1993b. Materialen zum Yemsa. Heidelberg, Universitätsverlag C. Winter.
- Lemma Didana. 1992 EC. Wolaytigna-Amharic English dictionary. Birhanuna Selam.
- Leslau, Wolf. 1959. *A dictionary of Moča (Southwestern Ethiopia)*. Berkeley and Los Angeles, University of California Press.
- Maghway, J. B. 1995. *Annotated Iraqw lexicon*. Tokyo, Research Institute for Languages and Cultures of Asia and Africa.
- Moseley, Christopher and R. E. Asher (eds.). 1994. *Atlas of the world's languages*. New York, Routledge.
- Parker, E. M. and R. J. Hayward. 1985. *An Afar-English-French dictionary (with grammatical notes in English)*. School of Oriental and African Studies, University of London.
- Reinisch, Leo. 1895. Wörterbuch der Bedauye-Sprache. Wien, Alfred Hölder.
- Savà, Graziano. 2005. A grammar of Ts'amakko. Köln, Rüdiger Köppe Verlag.
- Takahashi, Yona. [高橋洋成] 2009. "Hamarugo no kisogoi narabini dōshikeitai no kōsatsu [ハマル語の基礎語彙、ならびに動詞形態の考察]" [Basic lexicon and some notes on the verbal system of Hamar]. Hideyuki Inui (ed.). *Cushitic-Omotic studies 2008*, 107–138. Yamaguchi.
- Tilahun Gamta. 1989. Oromo-English dictionary. Addis Ababa University Printing Press.
- Tosco, Mauro. 1991. A grammatical sketch of Dahalo. Hamburg, Helmut Buske Verlag.
- Tosco, Mauro. 2001. The Dhaasanac language. Köln, Rüdiger Köppe Verlag.
- Tsuge, Yoichi [柘植洋一]. 1995. "Minami Omo gogun no on'intaiō ni tsuite. [南オモ語群の音韻対応について]" [On the phonetic correspondences of South Omotic languages]. *Tokyo University Linguistics Papers*, 14, 201–227.
- Tsuge, Yoichi. 2008. "A basic word list of the Dime language." Hideyuki Inui (ed.). *Cushitic-Omotic studies 2007*, 207–215. Yamaguchi.

- Wakasa, Motomichi [若狭基道]. 2000. "Woraitago goishū [ウォライタ語語彙集]" [A vocabulary of Wolaitta]. Tasaku Tsunoda (ed.). *Basic materials in minority language*, 65–112. Department of Asian and Pacific Linguistics, Institute of Cross-Cultural Studies, Graduate School of Humanities and Sociology, University of Tokyo.
- Wakasa, Motomichi [若狭基道]. 2016. "Woraitago no n on [ウォライタ語の n 音]" [The n sound in Wolaytta]. Presentation at the 53rd Annual Meeting of Japan Associations of African Studies.
- Wedekind, Klaus. 1990. "Gimo-Jan or Ben-Yem-Om: Benč_c Yemsa phonemes, tones, and words." Richard J. Hayward (ed.). *Omotic language studies*, 68–184. London, School of Oriental and African Studies, University of London.