

## **The Oromo Language (Oromic) Word Boundary**

### **Abstract**

Oromic speakers constitute the third largest number of speakers of the Afro-asiatic language group after Arabic and Hausa (Gragg, 1982), comprising about 35% of the more than 110 million population of Ethiopia alone (CIA World Fact book, 2021). Oromic speakers make up the largest national group in Ethiopia.

The language has been using Latin script based orthography since 1991. The pioneers who promulgated the orthography put down only the basic rules. With widespread use of the language for different purposes in the country, the need to update some rules has been felt. One of these rules is the demarcation of word boundaries indicated by spaces between the morphemes. The lack of these rules perpetuates some ambiguities.

Oromic has long words due to successive affixation morphology, both for inflection and derivation, as observed by Tutschek (1845). Yet, many of the morphemes that are affixed could be written disjunctively without causing any problems. These morphemes are mainly clitics that in many cases cause ambiguity when written conjoined with the host word. This is because some of the host morphemes have peripheral integral syllables homophonous with these clitics, making it difficult to tell whether they are part of the word or not. Besides, some of these unnecessary affixations include prefixing some clitics. Prefixes have been observed to be a burden when it comes to word recognition in the reading process (Lojenga, 2014a).

In this paper, morphemes that are indiscriminately affixed are listed and the ambiguities posed by writing these morphemes conjunctively also explained using the syntactic, phonological, and semantic criteria proposed by Lojenga (2012a).

Proponents of writing the morphemes conjoined argue that the basic rule of the orthography says we have to write as we speak, and we articulate these morphemes without any pause between them. Therefore, there is no need to separate them in writing either. As for the ambiguity mentioned above, they insist that native speakers can overcome most of these ambiguities via prosody or context. They further argue that these morphemes cannot stand alone because they will not make sense.

However, it does not seem that speakers pause at every word but mainly at phrase level, nor do we write all the morphemes we utter with one breath as a single word. Neither can we use prosody as the orthography has no provision for tone marking. Besides, with economic development more non-native people may be interested in learning the language. Additionally, there is a need to teach the language as a heritage language to children born to Oromo diaspora. Furthermore, digital machines need to be programmed for use, all of which require detailed rules to understand the meaning of the morpheme rather than depending on the context.

In conclusion, the Oromic orthography, being relatively new, will only benefit from continuous development of the rules. Word boundary rules help check the limitless length of the words and minimize, if not eliminate, ambiguities of arbitrary conjoining of morphemes. In this paper, the ambiguities associated with the lack of these rules, the morphemes implicated with them and the criteria to determine the word boundaries are addressed.

## Bibliography

- Andrzejewski, B.W. (1957). Some preliminary observations on the Borana dialect of [Oromo]. *Bulletin of the School of Oriental and African Studies*, 19(2), pp. 354-374. London: University of London.
- Banti, G. (1988). Two Cushitic systems: Somali and Oromo nouns. In van der Hulst, H. & Smith, N. (Eds.). *Autosegmental Studies on Pitch Accent*. Dordrecht: Foris Publications.
- Cahill, M., & Rice, K. (Eds.). (2014). *Developing orthographies for unwritten languages*. Dallas, TX: SIL, International.
- Carstairs, A. (1981). *Notes on Affixes, Clitics, and Paradigms*. Bloomington: Indiana University Linguistics Club.
- Clamons, Robbin, Ann E. Mulkern, Gerald Sanders, & Nancy Stenson. 1999. The limits of formal analysis: Pragmatic motivation in Oromo grammar. In Darnell, Michael, E. Moravcsik, F. Newmeyer, M. Noonan, & K. Wheatley (eds.), *Functionalism and Formalism in Linguistics: Case Studies*, 2, 59-76. Amsterdam: John Benjamins.
- Gasser, M. (2012). HornMorpho 2.5 user's guide. *Research Group: Human Language Technology and the Democratization of Information*. Indiana University.
- Gragg, G. B. (Ed.). (1982). *Oromo dictionary*. East Lansing, MI: African studies center, Michigan State University.
- Haspelmath, M., & Sims, A. D. (2010). *Understanding morphology* (2<sup>nd</sup> ed.). UK: Hodder Education.
- Haspelmath, M. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, 45(1), 31-80. doi 10.1515/flin.2011.002
- Janko, K. H. (2007). Causative verb and palatalization in Oromo: evidence from Oromo dialects and related Cushitic languages. *Journal of Oromo studies*, 14 (2), 89-104.
- Krapf, J. L. (1842). *Vocabulary of the Galla Language*. London: The Church Missionary Society.
- Krapf, J. L. (1860). *Travels, Researches, and Missionary Labours, During An Eighteen Years' Residence in Eastern Africa*. London: Trübner and Co., Paternoster Row.
- Lloret, M-R. (1995). The Representation of Glottals in Oromo *Phonology*, 12(2), 257-280. np: Cambridge University Press.

- Lloret, M-R. (1997). Oromo Phonology. In A. S. Kaye, *Phonologies of Asia and Africa*. (pp. 399-430). Winona Lake, IN: Eisenbrauns.
- Lojenga, C. K. (2014a). Basic principles for establishing word boundaries. In M. Cahill, & K. Rice, (Eds.), *Developing orthographies for unwritten languages* (pp. 73-106). Dallas, TX: SIL, International.
- Lojenga, C. K. (2014b). Orthography and tone: Tone-system typology with implications for orthography development. In M. Cahill, & K. Rice, (Eds.), *Developing orthographies for unwritten languages* (pp. 49-72). Dallas, TX: SIL, International.
- Mewis, C. J. (2001). *A grammatical sketch of written Oromo*. Koln: Rudiger Koppe Verlag .
- Owens, J. (1985). *A grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Buske.
- Rogers, H. (2005). *Writing systems: A linguistic approach*. MA: Blackwell Publishing.
- Schroeder, L. (2010). *Bantu Orthography Manual*. SIL e-Books 9 SIL International
- Sinha, S., & Talukdar, P.H. (2013). Word and syllable boundary of Sylheti phonemes/ syllables. *International Journal of Computing, Communications and Networking*, 2(1)
- US Central Intelligence Agency Website [https://www.cia.gov/library/publications/the-world-factbook/geos/print\\_et.html](https://www.cia.gov/library/publications/the-world-factbook/geos/print_et.html) Accessed Oct 23, 2020.
- Tucho, Y., Zorc, R.D., & Barna, E.C. (1996). *Oromo newspaper reader, grammar sketch and Lexicon*. Kensington, MD: dp Dunwoody Press.
- Tutschek, C. & Tutschek, L. (1845). *A grammar of the Galla language*. Munich: np.
- Zwicky, A. M., & Pullum, G. K. (1983). Cliticization vs. inflection: English N'T. *Language* 59(3), pp. 502-513.