

## Labial vibrants in Mangbetu

*Kenneth S. Olson* (SIL International and Laboratoire Parole et Langage) and  
*David Mbiri Koogibho* (Communauté Évangélique du Christ au Cœur d’Afrique)

### Abstract

This paper presents a digital list of words and short phrases that include labial vibrants (i.e. trills and flaps) in Mangbetu (D. R. Congo). The wordlist includes orthographic and broad phonetic transcriptions of each item, French and English glosses, MP3 and WAV recordings of each item, and metadata for resource discovery. An archival form of the list was deposited into an institutional archive (the Speech and Language Data Repository) which includes a WAV digital recording of the entire session, descriptive markup encoding of the list in XML employing Unicode transcription, and a metadata record. The presentation form was generated from the archival form.

### Zusammenfassung

Der vorliegende Artikel gibt eine digitalisierte Liste von Wörtern und kurzen Phrasen mit labialen Vibranten (Trills und Flaps) aus dem Mangbetu (Dem. Rep. Kongo). Jeder Eintrag der Wortliste ist in orthographischer und in breiter phonetischer Transkription wiedergegeben. Dazu kommen Audio-Aufzeichnungen in MP3 und WAV der einzelnen Einträge, Metadaten zur Herkunft. Eine Archiv-Version der Wortliste selbst liegt mitsamt der digitalen WAV-Aufzeichnung der vollständigen Aufnahme im Speech and Language Data Repository, verbunden mit einem Beschreibungs-Markup der Liste in XML (Unicode Transkription) und Metadaten. Die Darstellungsform wurde durch die Archiv-Version generiert.

### Introduction

- <1> This paper presents a digital list of words and short phrases in the Meegye variety of Mangbetu (ISO 639–3 code [mdj]), a Central Sudanic language spoken by perhaps 620,000 people in the northeastern part of the Democratic Republic of the Congo (Lewis, Simons & Fennig 2013).<sup>1</sup> The list exemplifies the occurrence of labial vibrants (i.e. trills and flaps) in the language.
- <2> This presentation form was generated from an archival form of the data (Olson et al. 2013). The procedure we followed for creating both forms is detailed in Simons, Olson & Frank (2007). In producing the archival form, we followed the best practice recommendations discussed in Bird & Simons (2003: 574–579) and Simons, Olson & Frank (2007).
- <3> In addition to a description of the primary data in the form of phonetic transcription, we provide a documentation of the data in the form of digital audio recordings (Himmelmann 1998), enabling the reader to verify and critique our transcription. This is important for these data, because Mangbetu has some unusual phonological characteristics.
- <4> First, Mangbetu has a three-way contrast between voiceless, voiced, and prenasalized bilabial trills [ɸ ɸ mɸ] (Larochette 1958, Tucker & Bryan 1966, Thomas, Bouquiaux & Cloarec-Heiss 1976, Demolin 1988, 1990, 1991, 1992, 2002, 2013, and McKee 1991,

---

<sup>1</sup> We wish to thank Mike Cahill, Didier Demolin, John Goldsmith, Larry Hyman, and Robert McKee for helpful comments. We take responsibility for all errors.

2007). Similar three-way contrasts have only been attested in the geographically nearby languages Asua, Lombi, Baka, and Lika (Keating 2007).

<5> The three bilabial trills in Mangbetu have been interpreted as simple phonemes (Larochette 1958, Demolin 1992) or as complexes consisting of a stop phoneme and a labial release (McKee 1991, 2007). That there are multiple possible phonemic interpretations in this case is a reflex of the non-uniqueness of phonemic solutions (Chao 1934). Regardless, all researchers agree that the sounds contrast with each other on the surface phonetic level. McKee (2007: 183) provides the following minimal triplet:

- |     |         |                            |
|-----|---------|----------------------------|
| (1) | [nóḅù]  | 'to bring out from within' |
|     | [nóḅù]  | 'to fan'                   |
|     | [nómḅù] | 'to enclose'               |

In addition, the three bilabial trills [ḅ ʙ mḅ] contrast with their respective bilabial plosives [p b mb] (Demolin 1992: 168,174, McKee 2007:183,185):

- |     |    |         |                  |                     |
|-----|----|---------|------------------|---------------------|
| (2) | a. | [nápú]  | 'ashes'          | #42 in our wordlist |
|     |    | [nóḅù]  | 'to push out'    | #14                 |
|     | b. | [nóbù]  | 'to feed a fire' | #43                 |
|     |    | [nábù]  | 'corpse'         | #1                  |
|     | c. | [némḅó] | 'have him tour'  |                     |
|     |    | [némḅó] | 'clay pot'       |                     |

<6> Second, bilabial trills are often thought to nearly always occur in a very specific environment, i.e. preceded by prenasalization and followed by a close back vocalic segment, and it is thought that they nearly always developed historically in this environment (e.g. Ladefoged & Maddieson 1996: 130). While this [mḅu] sequence is quite common cross-linguistically (Keating 2007), both Demolin and McKee provide robust evidence for additional environments for the realization of the sounds in Mangbetu: (a) the presence of both voiceless and plain (non-prenasalized) voiced bilabial trills in Mangbetu (cf. example 1 above) demonstrates that prenasalization does not obligatorily co-occur with a bilabial trill, and (b) the sample data in example (3) below show that the trills occur before vocalic segments other than close back ones in Mangbetu:

- |     |    |              |   |          |
|-----|----|--------------|---|----------|
| (3) | a. | [éḅì]        | 'leaping like a leopard'                    | #44, #59 |
|     |    | [éḅí]        | 'early morning'                             | #45, #60 |
|     |    | [néḅḅè]      | 'to blink'                                  | #50, #65 |
|     |    | [éḅándrè]    | 'toes'                                      | #52, #67 |
|     |    | [néìḅò]      | 'field'                                     | #9       |
|     | b. | [néḅàḅá]     | 'kind of plan'                              | #57, #72 |
|     |    | [néḅò]       | 'caterpillar'                               | #3       |
|     | c. | [émḅàmḅàmḅà] | 'water pulsing from a hose onto the ground' | #58, #73 |
|     |    | [nómḅò]      | 'to drink'                                  | #25      |

<7> In the cases of a bilabial trill followed by a front vowel, there is a noticeable labial glide at the intersection of the two segments (cf. Ladefoged, Cochran & Disner 1977: 54). Whether this glide should be interpreted as an integral part of the trill or as a distinct intervening segment has been the matter of some debate (Demolin 1992: 174, 2013; McKee 2007: 181–182). One factor to note is that there is no contrast between the presence or absence of the glide in this environment.

Note also that the presence of these additional environments is weak evidence against Ladefoged & Maddieson's claim that bilabial trills nearly always develop historically from an [mbu] sequence. An historical study of the Mangbetu case would be necessary in order to address that claim.

<8> Third, these three bilabial trills also contrast with a phonemic labiodental flap /ɽ/ in Mangbetu (Larochette 1958, Tucker & Bryan 1966, Demolin 1988, 1991, 1992, McKee 1991, Demolin & Teston 1996). This sound is sometimes realized as a bilabial flap in Mangbetu (Demolin 1992, Demolin & Teston 1996). Olson & Hajek (1999: 110, citing a personal communication from Robert McKee) provide the following minimal pair:

(4)	[nóvò]	'to defecate'	#35
	[nóbò]	'to get fat'	#6

<9> The presence of both labial trills and flaps in the same language is rare, and has only been attested in languages in geographic proximity to Mangbetu (Olson & Hajek 2003, Keating 2007).

Item 32 [névjàvjá] 'black bird' contributes one additional example of the rare occurrence of a labiodental flap in a consonant cluster (cf. Olson & Hajek 2003:168).

### Creating the wordlist

<10> The materials included in this presentation of the data include the following:

Wordlist: contains for each item a French and an English gloss, orthographic and broad phonetic transcriptions, and relevant notes.

1. Recordings: MP3<sup>2</sup> and WAV digital recordings of each item. The MP3 recording is accessible by clicking on the play icon next to the orthographic form of each word in the list below. Alternatively, you can play the WAV recording by clicking on the download icon. Your web browser will attempt to play the recording with the sound program that is set up as the default WAV player on your system. On Windows computers, you can download the WAV file by right-clicking on the download icon (control-click on Macintosh computers) and selecting "Save Target As..." or "Save Link As..." from the pull-down menu.
2. Metadata: a description of the data, useful for resource discovery.

<11> The original material consisted of an audio recording on two cassettes. The authors collected sample lexical items and short phrases, mostly from Demolin (1992) and McKee (2007) (the latter in draft form at the time). The second author transcribed the items in Meegye orthography (McKee 2002), adjusting them according to his own speech variety. Mangbetu exhibits much dialectal variation, so these adjustments should not be construed as invalidating the data from the other sources.

---

<sup>2</sup> Definitions of MP3, WAV, XML, XSLT, and other pertinent terms employed here are available at <http://emeld.org/school/glossary.html>.

- <12> The wordlist was recorded in March 2004 at the ACATBA center (l'Association Centrafricaine pour la Traduction de la Bible et d'Alphabétisation) in Bangui, Central African Republic. The recording was made with a Marantz PMD 420 monaural cassette recorder and an Audio-Technica ATM 33a condenser microphone. The second author, an adult male mother-tongue speaker of Meegye, produced the French gloss and one token of each item, reading from the orthographic transcription. Some items were then repeated a second time.
- <13> The recording was digitized at the International Linguistics Center (ILC) in Dallas, Texas in March 2005 by Roger E. Olson. The audio cassettes were played on a Marantz PMD 221 analog cassette recorder, and the recording was digitized onto a standard Windows XP computer using a Tascam US-122 USB Audio/Midi interface and CoolEdit 2000 for audio capture. The digitization was done at a 48 kHz sampling rate and a 24-bit quantization, which according to IASA-TC03 (2005:8) is the minimum recommended digital resolution for the archiving of analog originals. (Plichta & Kornbluh 2002 recommend 96 kHz and 24 bits.) The recordings were stored in non-compressed WAV format, also the recommended industry standard (IASA-TC03 2005:8).
- <14> At present, the hardware of some computers does not handle 48 kHz, 24-bit audio. For the purpose of the presentation form, then, the first author downsampled the audio recording using CoolEdit 2000 to 44.1 kHz, 16-bit (i.e. standard audio CD quality) applying a low-pass pre-filter to the data in order to prevent aliasing (Ladefoged 1996: 139–141). This sampling rate is sufficient for most technical purposes since it covers nearly all acoustic information pertinent to language (Ladefoged 2003: 18, 26). The higher quality original digital recording is available in the archival form.
- <15> Working from the orthographic version, the first author produced a broad phonetic transcription of the data which conforms to the extant version of the International Phonetic Alphabet (IPA 2006). This transcription is provided in the *Phonetic* column of the list. He also added clarifying notes to certain entries, found in the *Notes*<sup>3</sup> column.
- <16> He keyboarded the data into a Microsoft Word document using Unicode 5.1 characters.<sup>4</sup> He converted the data to a comma-delimited (CSV) file, and imported them into TableTrans v. 1.2 software (Bird et al. 2002). Judy Kuntz time-aligned the items to the original audio recording at the ILC in February 2008. This annotation was outputted to an XML<sup>5</sup> annotation graph output and transformed into an XML descriptive wordlist format using an XSLT script.
- The original digitized WAV file, the XML descriptive wordlist, and a metadata record constitute the archival form of the list. The metadata record follows the standard used by the Open Language Archives Community (OLAC).<sup>6</sup> The archival materials may be downloaded for free from the Internet (Olson et al. 2013).
- <17> The presentation form of the wordlist was then generated from the archival form. An XSLT script was employed to convert the archival XML descriptive wordlist into an HTML presentation wordlist. Then TableTrans was used to automatically create individual sound files corresponding to each of the segments identified in the transcription process for use in

---

<sup>3</sup> Abbreviations in the Notes column follow the Leipzig glossing rules, <http://www.eva.mpg.de/lingua/resources/glossing-rules.php>.

<sup>4</sup> <http://www.unicode.org>.

<sup>5</sup> <http://www.w3.org/XML/>.

<sup>6</sup> <http://www.language-archives.org/OLAC/metadata.html>. An alternative standard for linguistic metadata has been proposed called the ISLE Meta Data Initiative <http://www.mpi.nl/imdi/>. See also <http://www.clarin.eu/cmd/>.

the presentation form. In-browser playing of the MP3 audio files was implemented using UbaPlayer 1.0.0,7 an open source HTML5 audio player with Adobe Flash fallback.

## Discussion

- <18> Some best practice recommendations were not followed in the creation of this resource. First, the items are spoken in isolation rather than in carrier sentences. A carrier sentence may help avoid list intonation and make it easier to measure the length of some sounds (e.g. word-initial stops) (cf. Ladefoged 2003:7–8). Second, only one speaker was recorded. Recording a larger number of speakers could help ensure that the data reflect the language as a whole and not an individual's idiolect. Ladefoged (2003:14) recommends recording half a dozen speakers of each sex. Third, only one token of several items was recorded. Fourth, the original hand-written orthographic transcriptions were not retained, and hence not archived.
- <19> On the other hand, the project succeeded in following most of the best practice recommendations found in Bird & Simons (2003:574–579) and Simons, Olson & Frank (2007). For example, the primary recording was digitized at industry-standard archival rates and stored in open (i.e. published) formats. The transcriptions were made in both IPA and orthography, encoded in Unicode, time-aligned to the underlying recording, and stored in an open descriptive markup format. The archival form was deposited into an institutional archive dedicated to the long-term preservation and availability of the resource. A human-readable version of the resource using presentational markup was provided (i.e. this web resource). Finally, the resource was described using metadata listed with an industry-standard repository for resource discovery.

## References

- Bird, Steven, Kazuaki Maeda, Xiaoyi Ma, Haejoong Lee, Beth Randall and Salim Zayat 2002 'TableTrans, MultiTrans, InterTrans and TreeTrans: Diverse tools built on the Annotation Graph Toolkit.' In: *Proceedings of the Third International Conference on Language Resources and Evaluation*. Paris: European Language Resources Association. <http://arxiv.org/abs/cs/0204006> (10.05.2013)
- Bird, Steven and Gary F. Simons 2003 'Seven dimensions of portability for language documentation and description.' *Language* 79,3: 557–582. <http://dx.doi.org/10.1353/lan.2003.0149> (05.10.2013). Preprint: <http://arxiv.org/abs/cs/0204020> (10.05.2013)
- Chao, Y. R. 1934 'The non-uniqueness of phonemic solutions of phonetic systems.' *Bulletin of the Institute of History and Philology, Academia Sinica* 4: 4, 363–397
- Demolin, Didier 1988 'Some problems of phonological reconstruction in Central Sudanic.' *Belgian Journal of Linguistics* 3:53–95. <http://dx.doi.org/10.1075/bjl.3.06dem> (10.05.2013)
- Demolin, Didier 1990 'Les trilles bilabiales du mangbetu.' *Pholia* 5: 67–90. [http://www.ddl.ish-lyon.cnrs.fr/Download/Pholia/Pholia\\_N-5.pdf](http://www.ddl.ish-lyon.cnrs.fr/Download/Pholia/Pholia_N-5.pdf) (10.05.2013)

---

<sup>7</sup> <http://www.brianhadaway.com/ubaplayer-the-jquery-html5-audio-player-with-flash-fallback/>.

- Demolin, Didier 1991  
'L'analyse des segments, de la syllabe et des tons dans un jeu de langage Mangbetu.'  
*Langages* 25,101: 30–50. <http://dx.doi.org/10.3406/lgge.1991.1800> (10.05.2013)
- Demolin, Didier 1992  
Le mangbetu: Étude phonétique et phonologique. Unpublished Ph. D. Diss. Brussels:  
Université Libre de Bruxelles
- Demolin, Didier 2002  
'The search for primitives in phonology and the explanation of sound patterns: The contribution of fieldwork studies.' In: Carlos Gussenhoven & Natasha Warner (eds.),  
*Laboratory phonology* 7, pp. 455–513. (Phonology and Phonetics, 4–1.) Berlin: Mouton de Gruyter
- Demolin, Didier 2013  
The phonetics of bilabial trills in Mangbetu. Paper presented at the conference  
"Phonetics and Phonology of Sub-Saharan Languages", Johannesburg, 7–10 July 2011
- Demolin, Didier and Bernard Teston 1996  
'Labiodental flaps in Mangbetu.' *Journal of the International Phonetic Association*  
26, 2: 103–111, <http://dx.doi.org/10.1017/S0025100300006149> (10.05.2013)
- Himmelman, Nikolaus 1998  
'Documentary and descriptive linguistics.' *Linguistics* 36, 1: 161–195.  
<http://dx.doi.org/10.1515/ling.1998.36.1.161> (05.10.2013), full version:  
<http://www.hrelp.org/events/workshops/eldp2005/reading/himmelman.pdf>  
(10.05.2013)
- IASA-TC03 2005  
'The safeguarding of the audio heritage: Ethics, principles and preservation strategy,  
version 3.' <http://www.iasa-web.org/tc03/ethics-principles-preservation-strategy>  
(10.05.2013)
- International Phonetic Association 2006  
'IPA news.' *Journal of the International Phonetic Association* 36,1: 133–135.  
<http://dx.doi.org/10.1017/S0025100306002465> (10.05.2013)
- Keating, Christine A. 2007  
A cross-linguistic overview of bilabial trills. Unpublished M.A. Thesis Dallas: Graduate  
Institute of Applied Linguistics
- Ladefoged, Peter 1996  
*Elements of acoustic phonetics*. Second edition. Chicago: University of Chicago Press
- Ladefoged, Peter 2003  
*Phonetic data analysis: An introduction to fieldwork and instrumental techniques*.  
Oxford: Blackwell
- Ladefoged, Peter, Anne Cochran and Sandra Disner 1977  
'Laterals and trills.' *Journal of the International Phonetic Association* 7,2: 46–54.  
<http://dx.doi.org/10.1017/S0025100300005636> (04.09.2013)
- Ladefoged, Peter and Ian Maddieson 1996  
*The sounds of the world's languages*. Oxford: Blackwell
- Larochette, J. 1958  
*Grammaire des dialectes mangbetu et medje suivi d'un manuel de conversation et d'un  
lexique*. (Annales du Musée Royal du Congo Belge 8. Sciences de l'Homme 18).  
Tervuren: Musée Royal du Congo Belge

- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds.) 2013  
*Ethnologue: Languages of the world*. Seventeenth edition. Dallas, TX: SIL International. <http://www.ethnologue.com> (10.05.2013)
- McKee, Robert G. 1991  
 'The interpretation of consonants with semi-vowel release in Meje (Zaire) stems.' In: Franz Rottland & Lucia N. Omondi (eds.), *Proceedings of the Third Nilo-Saharan Linguistics Colloquium, Kisumu, Kenya, August 4–9, 1986*, pp. 181–195. Hamburg: Helmut Buske
- McKee, Robert G. 2002  
 Mangbetu orthography statement, first revision. Bunia: SIL–Eastern Congo Group. ms.
- McKee, Robert Guy 2007  
 'Concerning Meegye and Mangbetu's bilabial trills.' In Doris L. Payne & Mechthild Reh (eds.), *Advances in Nilo-Saharan Linguistics. Proceedings of the 8th Nilo-Saharan Linguistics Colloquium, University of Hamburg, August 22–25 2001*, 181–189. (Nilo-Saharan. Linguistic Analyses and Documentation 22). Cologne: Rüdiger Köppe
- Olson, Kenneth S. and John Hajek 1999  
 'The phonetic status of the labial flap.' *Journal of the International Phonetic Association* 29,2:101–114. <http://dx.doi.org/10.1017/S0025100300006484> (10.05.2013)
- Olson, Kenneth S. and John Hajek 2003  
 'Crosslinguistic insights on the labial flap.' *Linguistic Typology* 7,2:157–186. <http://dx.doi.org/10.1515/lity.2003.014> (10.05.2013)
- Olson, Kenneth S. and John Hajek 2004  
 'A crosslinguistic lexicon of the labial flap.' *Linguistic Discovery* 2,2: 21–57. <http://linguistic-discovery.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/article/262> (10.05.2013)
- Olson, Kenneth S., David Mbiri Koogibho, Roger E. Olson, Judy Kuntz and Gary F. Simons 2013  
 'Labial vibrants in Mangbetu: Archival form.' Primary data (corpus). Speech and Language Data Repository. <http://hdl.handle.net/11041/sldr000829> (10.06.2013)
- Plichta, Bartek and Mark Kornbluh 2002  
 'Digitizing speech recordings for archival purposes.' [http://www.historicalvoices.org/papers/audio\\_digitization.pdf](http://www.historicalvoices.org/papers/audio_digitization.pdf) (10.05.2013)
- Simons, Gary F., Kenneth S. Olson and Paul Frank 2007  
 'Ngbugu digital wordlist: A test case for best practices in archiving and presenting language documentation.' *Linguistic Discovery* 5,1: 28–39. <http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/article/314> (10.05.2013)
- Thomas, Jacqueline M.-C., Luc Bouquiaux and France Cloarec-Heiss 1976  
*Initiation à la phonétique: Phonétique articulatoire et phonétique distinctive*. [Paris]: Presses Universitaires de France
- Tucker, A. N. and M. A. Bryan 1966  
*Linguistic analyses: The non-Bantu languages of North-Eastern Africa*. London: Oxford University Press (for the International African Institute)