Perception of Yoruba Word-Initial [g͡b] and [b]

Michael Cahill
SIL International

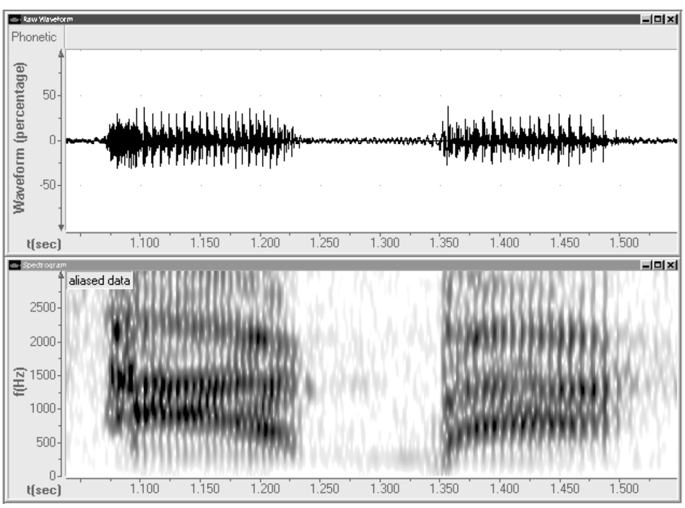
1. Introduction

Labial-velar obstruents /g͡b/, /k͡p/, /ŋ͡m/ occur in hundreds of African languages in sub-Saharan Africa, dozens of languages of the Pacific (mostly in Papua New Guinea and Vanuatu), and a handful of languages in South America (Cahill 1999 and references therein). To date 615 languages in my personal database have phonemic labial-velars, approximately 8.9% of the 6912 languages in the most recent Ethnologue (Gordon 2005), and I expect a more thorough count will increase this figure to approximately 10% of the world's languages.

This paper is a preliminary investigation into what cues are used by native Yoruba speakers to distinguish phonetic [g͡b] and [b]. More specifically, I focus on word-initial occurrences of these sounds, not on intervocalic position (e.g. [ag͡ba]) or post-nasal position (e.g. [ŋg͡ba]). One might assume that word-initial occurrences (e.g. [g͡ba] vs. [ba]) would be inherently more difficult to distinguish than intervocalic occurrences (e.g. [ag͡ba] vs. [aba]), for at least two reasons. First, both CV and VC cues are present in [ag͡ba], while only CV transition cues are present in [g͡ba].

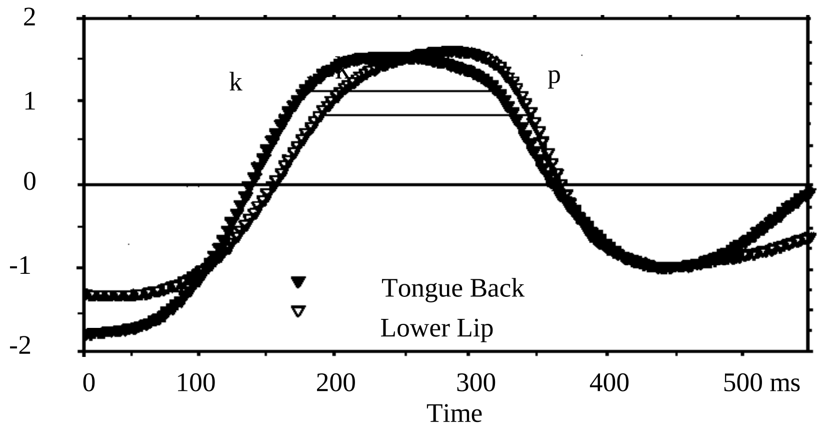
Second, though labial-velars are commonly described as having "simultaneous" labial and velar articulations, this is not quite accurate. They have overlapping but not coterminous articulations. The start of the velar articulation precedes the labial and ends sooner as well. Thus, the release of a labial-velar is labial (Ladefoged and Maddieson 1996 and references therein). This is shown in formant transitions in spectrograms from several languages in Connell 1994 and Ladefoged and Maddieson 1996, as well as below (Note especially the difference in F2 in transitions into and out of the [gb].)

(1) Spectrogram of *agba* from Leggbo (Nigeria - recording courtesy of Julie Larson)



It is also shown more directly by means of electromagnetic articulography, in which small metal pellets are glued to various positions on the tongue and lips, and sensors record the position of these over time, giving data such as below.

(2) Electromagnetic articulography shows overlap of the gestures (Ladefoged and Maddieson 1996).



Coordination of lower lip and tongue back movements in the Ewe word *ak͡pa*. Y-axis is vertical displacement; horizontal lines indicate the likely duration of actual contact of the articulator.

Since both labial-velars and labials have a labial release, the CV formant transitions of both labial-velars and labials would be expected to be similar. (Actually, we will see that there is a difference.)

Though word-initial [g͡b] vs. [b] are thus expected to be difficult to distinguish, and indeed to untrained Western ears they sound identical, native speakers, including Yorubas, reliably and consistently distinguish these. The question investigated here is what factors cue the identification of [g͡b] vs. [b] in Yoruba. (Yoruba has no voiceless [k͡p] vs. [p] contrast.) Two factors were investigated: first, CV transition cues and second, prevoicing information. Results show that each of these has a role in cuing Yoruba speakers to reliably distinguish word-initial [g͡b] and [b].

2. Methodology

All recordings and file manipulations, with a single exception mentioned below, were done with SIL's Speech Analyzer software.

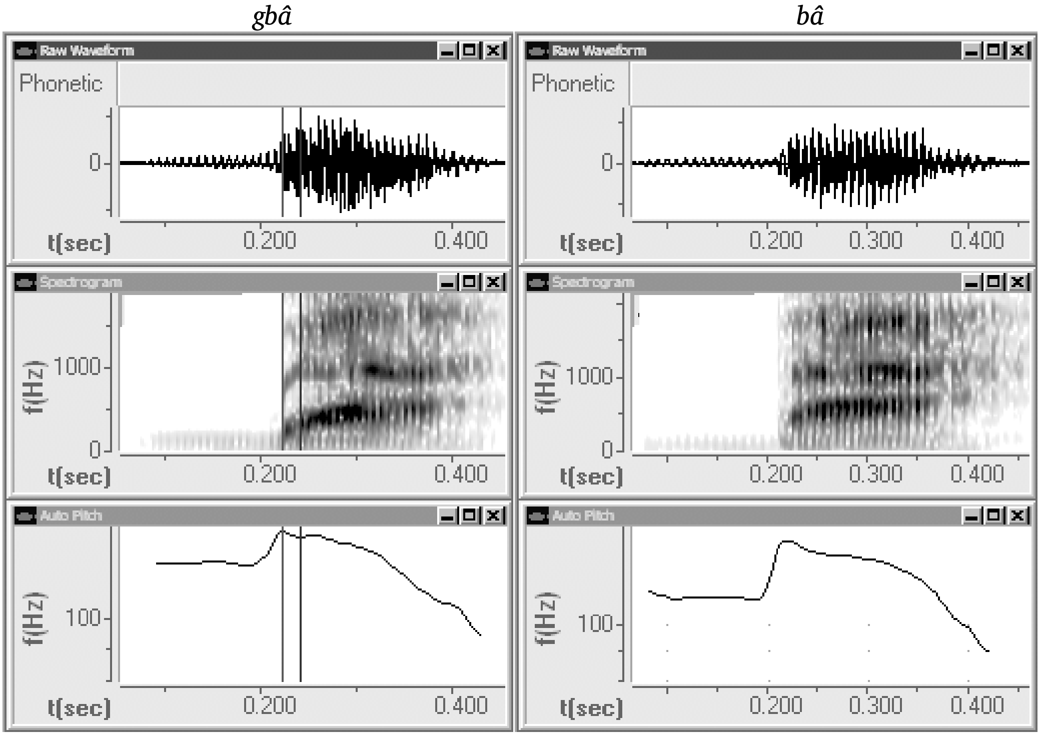
A native Yoruba speaker and I perused a Yoruba dictionary together and selected eight minimal pairs differing only in the [g͡b] ~ [b] parameter (e.g. *g͡bɔ́ /bɔ́*, *g͡bɛ̀rɛ̀ /bɛ̀rɛ̀*). He recorded these directly into the computer, with three repetitions for each utterance. The second of the three repetitions was selected for further processing, detailed below. Since two recordings were slightly quieter than the others, the intensity of these was boosted with the CoolEdit program, so that all words had approximately the same volume.

Randomized sets of recordings in the form of computer sound files were played to 5 native Yoruba speakers. The speakers were all fluent speakers living in the Dallas area, with ages from 21-50 years old, 2 females and 3 males, who identified the token as either a [g͡b]-initial word or a [b]-initial word (e.g. *ba* or *gba*) by circling their choice. Listeners heard each token twice, and had a chance to ask for a repeat, which was rarely taken.

- The first set of sound files had unmodified recordings of the minimal pairs.
- In the second set, the prevoicing was excised from the [g͡b] word and replaced the prevoicing in the corresponding [b] word, and vice versa, i.e. prevoicings were switched (see waveforms in (3)). Note that the durations of prevoicings were the same.

- In the third set, the steep portion of the F2 of the CV transition of [g͡b] (averaging 18 ms) was cut from the [g͡b] words and inserted into the [b] words (see the position of the cursors in [g͡ba] in (3)).

(3) Waveform, spectrogram, and pitch trace for Yoruba ***gbâ*** and ***bâ***:



Finally, the last set was repeated to check for consistency and subject fatigue. So sound file Set 4 is identical to Set 3.

All cuts and insertions were made at zero crossings to minimize pops (an optional setting of Speech Analyzer). Each set had 24 tokens, so the test consisted of listening to and identifying 96 total tokens, which typically took a total of 15 minutes.

3. Results

With the unmodified recordings, there was almost 100% correct identification of the words. Four of five subjects identified 100% of tokens correctly, while one subject missed one token. In contrast, in Set 2, where the prevoicings had been switched, subjects identified 75% of tokens with their original bases. In Sets 3 and 4, having tokens with the added or deleted CV transitions, subjects identified only 56% of the tokens with their original bases.

Consistency of identification was, as mentioned, virtually unanimous in the unmodified recordings; every one agreed with each other. In contrast, there were only a handful of the modified recordings in which all subjects agreed. (8/24 in Set 2, 5/24 in Set 3, 4/24 in Set 4.) As mentioned, Sets 3-4 were identical, repeated to test for subject consistency. In these, subjects agreed with themselves 88% of the time, i.e. consistency was fairly high.

A result also examined was directionality: was there any tendency for subjects to identify more tokens as *g͡b* rather than *b*, or the reverse? It turns out that there was a slight tendency to identify more tokens as *g͡b* when it was originally *b* than the reverse, but this was <u>highly</u> subject-dependent, as can be seen from the chart below. Subjects 1, 3, and 5 had a bias toward identifying tokens as *g͡b*, while Subjects 2 and 4 had a bias the other way, though Subject 4 was not as extreme as Subject 2.

(4) Identification of tokens that were modified

All changed tokens	original ***g͡b*** identified as ***b***	original ***b*** identified as ***g͡b***	TOTAL “missed”
Subject 1	2	25	27
Subject 2	22	8	30
Subject 3	3	19	22
Subject 4	20	10	30
Subject 5	9	17	26
TOTALS	56	79	135 *(out of 360 tokens)*

The conclusion to be drawn is that both pre-voicing and the CV release are factors in distinguishing /*g͡b*/ from /b/ in Yoruba, with the CV release being somewhat more important.

4. Discussion

Let us examine some details of the prevoicing factor, then the CV release.

Prevoicing has previously been identified as a cue for identifying certain sounds (e.g. Kutsch Lojenga 1991 and Demolin 1995 for “voiceless implosives”), but this has not been a common topic for research and has not previously been asserted as a cue for identifying labial-velars. Nonetheless, the results above show a clear influence of prevoicing as a factor in distinguishing *g͡b* and *b* in Yoruba. Several acoustic differences between the prevoicings of [g͡b] and [b] were observed which may account for these experimental results.

The amplitude of the prevoicing often (but not always) increased during the period of [g͡b] closure. This increasing amplitude can be seen in the waveform in (3) for *g͡b*, while the prevoicing for *b* is steadier, and actually decreases slightly. This increase in amplitude of the prevoicing for *g͡b* is consistent with an implosive air mechanism, that is, a more detailed transcription would be more like [g͡ɓ]. (See Lindau 1984).

Another phonetic detail of the prevoicing had to do with pitch. The F0 in the prevoicing of [g͡b] was consistently higher than that of [b], as seen in the pitch tracing in (3). I speculate that this is may be an enhancement feature of implosives. Most, perhaps all, research on pitch and implosives has focused on F0 of the following *vowel*, not of the internal pitch of the implosive itself. Supporting the hypothesis of higher pitch as an indicator of implosive and not of labial-velars *per se* is data from Kɔnni, a Gur language of northern Ghana, in which labial-velars are non-implosive, e.g. ***g͡bágʊ́bá*** ‘forest squirrel.’ In Kɔnni, the pitch of the prevoicing of both [g͡b] and [b] is comparably low. (Of course, since this removes a possible cue, the question arises what cues Kɔnni speakers use to distinguish word-initial *g͡b* and *b*.)

Let us now turn to the CV release details. Connell (1994) and others have noted that the CV formant transitions for the labial-velars resemble those for simple labials but are steeper, having a lower locus, and possibly more intense, presumably cuing labial-velar recognition. However, again there has been no previous perceptual study of the question. Again, the spectrograms in (3) show that the CV transition of F2 of [g͡b] in Yoruba is steeper, with a lower locus, than the [b]. This is perhaps surprising in light of the fact that labial-velars have a labial release, and one would expect the CV transitions of these to be more alike, if not identical. But they are distinctly different. Since Connell and others made their study with a variety of languages, with both implosive and non-implosive labial-velars (and I have seen the same CV release pattern in Kɔnni), it seems likely that the distinctive release characteristics of labial-velars are not a result of an implosive air mechanism, but are a result of some as-yet unknown phonetic factors specific to labial-velars.

The results demonstrate that there is information both in the pre-voicing and in the CV transition which contributes to identification of a consonant as [g͡b] vs. [b].

Future investigations are called for, of course. The most obvious is the need to perform more statistical analysis on the data in this study. It would also be desirable to record other Yoruba speakers and confirm the characteristics of prevoicing noted here. Another more stringent Yoruba study might study minimal pairs in a more natural language context, without the opportunity for repeats. Other languages, in which [g͡b] is not implosive, as well as those in which it is, need to have similar studies made on them, as well as languages which contrast implosive [g͡b] and [ɓ]. Finally, the labial-velars /k͡p/ and /ŋ͡m/ do not have pre-voicing, but have other cues, and these need to be explored as well.

References

Cahill, Michael. 1999. Aspects of the phonology of labial-velar stops. Studies in African Languages 28.155-184.

Connell, Bruce. 1994. The structure of labial-velar stops. Journal of Phonetics 22.441-476.

Demolin, Didier. 1995. The phonetics and phonology of glottalized consonants in Lendu. Phonology and Phonetic Evidence. Papers in Laboratory Phonology IV, ed. by Bruce Connell and Amalia Arvanti, 368-385. Cambridge: Cambridge University Press.

Gordon, Raymond G., Jr. (ed.) 2005. Ethnologue: Languages of the World, Fifteenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com/.

Kutsch Lojenga, Constance. 1991. Lendu: a new perspective on implosives and glottalized consonants. Afrika und Übersee 74.77-86.

Ladefoged, Peter, and Ian Maddieson. 1996. The Sounds of the World’s Languages. Cambridge, MA: Blackwell Publishers.

Lindau, Mona. 1984. Phonetic differences in glottalic consonants. Journal of Phonetics 12.147-155.

Maddieson, Ian. 2005. Presence of uncommon consonants. World Atlas of Language Structures, ed. by Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, 82-85. Oxford: Oxford University Press.

Selected Proceedings of the 36th Annual Conference on African Linguistics: Shifting the Center of Africanism in Language Politics and Economic Globalization

edited by Olaoba F. Arasanyin and Michael A. Pemberton

Cascadilla Proceedings Project Somerville, MA 2006

Copyright information

ISBN 1-57473-414-8 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Cahill, Michael. 2006. Perception of Yoruba Word-Initial [g͡b] and [b]. In *Selected Proceedings of the 36th Annual Conference on African Linguistics*, ed. Olaoba F. Arasanyin and Michael A. Pemberton, 37-41. Somerville, MA: Cascadilla Proceedings Project.

or:

Cahill, Michael. 2006. Perception of Yoruba Word-Initial [g͡b] and [b]. In *Selected Proceedings of the 36th Annual Conference on African Linguistics*, ed. Olaoba F. Arasanyin and Michael A. Pemberton, 37-41. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1404.