VOICING OF LABIOVELAR STOPS IN YORUBA

Sven Grawunder[a], *Bodo Winter*[a,b] *& Joseph Atoyebi*[a]

[a]Department of Linguistics, Max Planck Institute for Evolutionary Anthropology, Germany;
[b]Department of Cognitive and Information Sciences, University of California, Merced, USA
grawunder@eva.mpg.de; bodo@bodowinter.com; atoyebi@eva.mpg.de

ABSTRACT

This production study investigates the correlates of voicing in initial /b, k, g, g͡b, k͡p/ stops in Yoruba. We found “voiceless” /k͡p/ to be “partially voiced”. It has prevoicing, but at the same time it behaves like a typical voiceless stop in raising the f0 of the following vowel, and in having the tendency to decrease surrounding vowel durations. /k͡p/ and /g͡b/ are furthermore distinguished from each other with respect to the duration of prevoicing. The single stops /k/ and /g/, on the other hand, differ not only with respect to voice onset time, f0 and vowel duration, but also with respect to the complete absence or presence of prevoicing.

We furthermore focus on looking for voicing enhancement mechanisms in the stop voicing system. Preliminary analysis of EGG data and the intensity during the prevoicing seems to suggest that for voiced labiovelar stops, there might be a voicing enhancement mechanism that is stronger than in the case of voiced single stops.

Keywords: labiovelar double articulation, voicing constraints, prevoicing, EGG

1. INTRODUCTION

Labiovelar double articulations are quite rare in the world’s languages, but they appear frequently in the languages of Central and Subsaharan West Africa [5, 8, 12]. Labiovelar stops have been subject to a number of phonetic studies in different languages of the area, where their aerodynamic characteristics [8, 12], the timing of the two gestures [13] or the acoustics of the labiovelar release have been investigated [6, 11].

The current paper is part of a larger production study that aims at a complete description of the phonetics of initial Yoruba labiovelar double articulations – incorporating acoustic recordings, EGG, airflow measurements, automated visual lip tracking and ultrasound. For this paper, we focus on characterizing the voicing properties of the Yoruba stops.

Yoruba is a language that has both “voiced” /g͡b/ and “voiceless” /k͡p/ [1, 12]. This allows us to compare the voicing contrast of labiovelar double articulations to the voicing contrast of singly articulated stops, although only for /b, k, g/ since /p/ does not exist in Yoruba. Specifically, we concentrate on the *initial* voicing contrast because VC transitions are not available as cues in this position and therefore, other cues should come into play.

In some languages, /k͡p/ appears to be voiced [8] – despite being described as “voiceless”, as is the case with phonological descriptions of Yoruba [1, 14, 17]. Given that /k͡p/ and /g͡b/ share the same places of articulation, the question arises as to what distinguishes these segments from each other if they are actually both “voiced”.

Another reason to investigate the labiovelar double articulations in Yoruba is that these stops tend to be realized with non-pulmonic (velaric) airstream mechanisms [13] and often show an implosive component [8, 10, 12] p. 44. If this is really the case, how can a voicing contrast between /g͡b/ and /k͡p/ be realized? What are the correlates of voicing, and do they differ from the correlates found in single articulations?

2. METHODOLOGY

2.1. Speakers, stimuli and procedure

Five native speakers of Yoruba (3 males, 2 females) were recorded. All speakers reported to have acquired Yoruba from their parents, and they reported to use the language frequently despite living in Leipzig (Germany).

We constructed a stimuli list of 75 monosyllabic words. The segments /b, k, g, g͡b, k͡p/ were combined with the vowels /a, e, ɛ, i, ĩ, o, ɔ, õ, ɔ̃, u/ and the three tones L, M and H (see Table 1). Representative words are *gbọ́* ‘to hear’, *pò* ‘to mix’, *gɔ* ‘to hide’ and *bɔ́* ‘to feed’.

There were five blocks; the stimuli order was randomized within blocks. In all blocks, we recorded acoustics and EGG. In blocks 3-5, we additionally recorded airflow.

Table 1: Overview of stimuli properties.

Segments	No.	Vowels	No.
gb	14	a	11
kp	10	e/ɛ	16
g	14	i	8
k	19	ĩ	7
b	18	o/ɔ	22
Tones	No.	õ/ɔ̃	6
L	27	u	5
M	19		
H	29		

For each trial, participants were asked to first read the word in isolation and then in the carrier phrase *tún pe ọ̀rọ̀ yìí* ___ 'Repeat the word ___'.

2.2. Recordings and acoustical analysis

Only the acoustic and the EGG recordings will be analyzed in this paper. The recordings were carried out in the sound-proof booth of the MPI EVA phonetics lab. A Microphone (*Sennheiser M60 + K6)* was placed approximately 15 cm in front of the speakers. For laryngography, we used an *EG-2 Glottal Enterprises* double-channel electroglottograph with electrode gel *spectra 360* and consistent angular positioning of the electrodes for all subjects. All signals were recorded via the digital multi-track recorder *Sound Devices 788T* in 48kHz/24bit.

All acoustic materials were checked for mispronunciations and manually annotated. The first visible zero crossing of glottal fold vibration was taken as the beginning of prevoicing. A rapid energy build-up and a concomitant burst- or aspiration-like noise was taken to be the beginning of the stop release. The onset of the vowel was defined as the zero crossing of the first period for which F2 was visible.

We collected the following dependent measures: (1) prevoicing duration, (2) release-to-vowel-onset duration, (3) vowel duration, (4) vowel f0 and (5) prevoicing intensity.

All data were analyzed using *R* and linear mixed effects models with the packages *lme4* [3] and *languageR* [2]. For each dependent measure, we first constructed a general (="omnibus") model with the fixed effects "Stop" (5 levels: /b, k, g, g͡b, k͡p), "Repetition" (1 to 5), "Context" (isolation vs. carrier phrase), "Tone" (L vs. M vs. H) and the random effects "Subject" and "Item". If the factor "Stop" reached significance, we performed individual comparisons (e.g. "/g͡b/ vs. /k͡p/") with linear mixed effects models. These comparisons do not need to be Bonferroni-corrected because the omnibus test already refutes the global null hypothesis (= the family-wise error rate is controlled for). Throughout the paper, we report MCMC-estimated p-values of validated models (likelihood-ratio test of null model against test model). In case the data did not meet the normality or homogeneity requirements, we performed data cleaning (2SD cut-offs) or transformations (e.g. square root).

3. RESULTS

3.1. Prevoicing and prevoicing duration

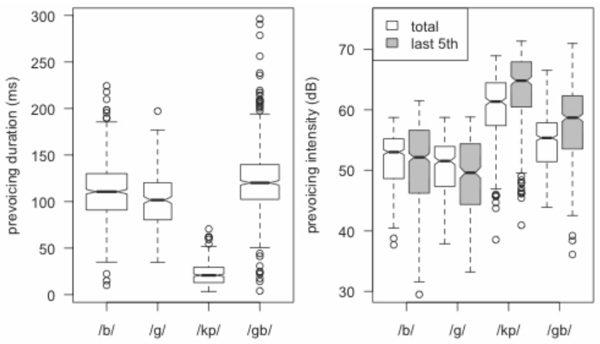
On average, Yoruba voiced stops have substantially long prevoicing durations (111 ms). The "voiceless" /k͡p/ segments are almost always realized with prevoicing (95%), however, the prevoicing duration is much shorter than in the case of voiced stops (19 ms). The prevoicing of /k͡p/ is both shorter in comparison to all of the other stops ($p<0.0001$) and in comparison to /g͡b/ ($p<0.0001$), which shows prevoicing in 99% of all words. Thus, the duration of prevoicing is one feature that distinguishes /k͡p/ from /g͡b/. Other contrasts, in particular /b/ vs. /g/, /g/ vs. /gb/, and /b/ vs. /g͡b/, exhibit no overall significant effects with respect to prevoicing duration (see Fig. 1).

3.2. Prevoicing intensity

We believe that the intensity and the duration of prevoicing can be used as a shorthand to assess the "ease" of glottal fold vibration before the release. For example, for /g/, the pressure builds up relatively quickly during phonation due to the small cavity size, rendering voicing difficult [15]. The intensity of glottal fold vibration, as well as the duration through which prevoicing can be sustained, might be correlates of this voicing constraint. Moreover, during initial screenings of the data, we noticed that the development of prevoicing intensity (as can be seen from intensity curves) seemed to be different for different stops. We therefore measured the intensity in 5 successive intervals during prevoicing, expecting the last interval (at point 5 just before the release) to show the biggest effect of the voicing constraint due to the largest pressure.

Fig.1 shows the prevoicing intensity for /g/ and /g͡b/. Both stops have velar closures and should thus follow the same voicing constraint, however, /g͡b/ has higher intensities (+4dB) than /g/. This difference is not significant when average intensities are considered (p=0.22), but it is significant at point 5 just before the release (p= 0.015) when the influence of the voicing constraint is expected to be strongest. From this perspective, the fact that there is a difference between /g͡b/ and /g/ might point towards a possible voicing enhancement mechanism (e.g. cavity extension through lowering of the larynx).

Figure 1: Prevoicing duration (left) and prevoicing intensity (right).



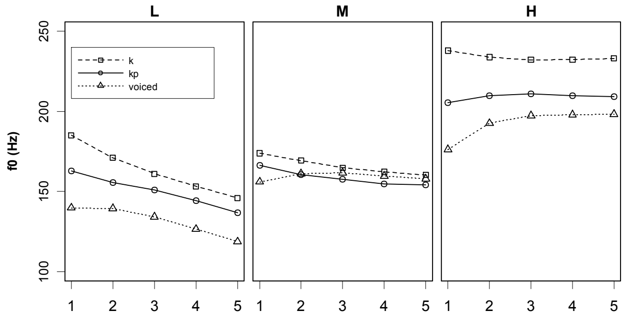
3.3. F0 and duration of the following vowels

Given that both /k͡p/ and /g͡b/ have prevoicing, we looked for other possible cues of the voicing contrast. We found that f0 is higher for /k͡p/ than for /g͡b/. As average across the 5 measure points, this difference is only marginally significant (p=0.0790), but at vowel onset, where the micro-prosodic influence of the stop on vowel pitch is expected to be strongest, the effect is significant (p=0.0140). Fig. 2 shows (for male speakers only) the vowel-internal f0 development following the stop consonant (female speakers exhibit a very similar pattern).

In general, phonemically "voiceless" stops (/k/ and /k͡p/) lead to a higher f0 both on average (p=0.0001) and at vowel onset (p=0.0001) compared to voiced stops. Voiceless segments were on average about one semitone higher, a difference that is well within the JND range of pitch. Moreover, even though they behaved quite similarly, there was a discernable difference between /k/ and /k͡p/ in that /k/ lead to a more expressed f0 difference (on average: p=0.0302, at vowel onset: p=0.0056).

With respect to vowel duration, the vowels following /k/ and /k͡p/ were shorter than following the voiced stops (p=0.068). With regards to this measurement, /k/ and /k͡p/ actually behaved statistically indistinguishable from each other (p=0.64).

Figure 2: Tone and f0 in the following vowel for voiced vs. voiceless labiovelar stops of the three male subjects; the numbers on the x-axis refer to the measurement steps.



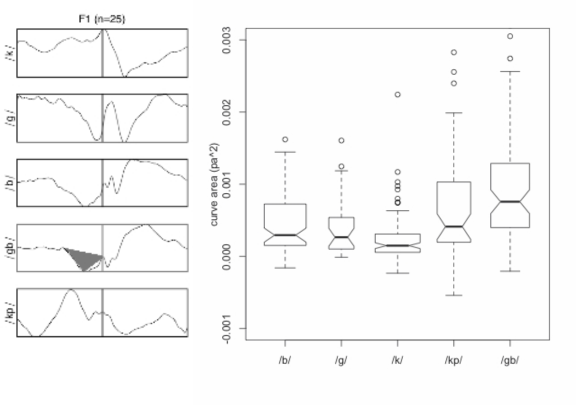
3.4. EGG – Gx movement

In order to assess the possibility that there is a voicing enhancement mechanism at play, we looked at 2-channel EGG signals. Precise derivation of larynx height is admittedly difficult since there is a multitude of factors influencing the impedance between the electrodes, but global D/C behavior (Gx) might be taken as a rough indicator for larynx displacement *if it occurs systematically with specific stops*. Therefore we operate with the rationale that wide-ranging synchronous parallel vertical D/C is an indicator of large tissue displacement across both electrodes, vis-à-vis narrow symmetric Gx movements which can be indicators for tissue displacement between the two electrode parts. Therefore, both channels were merged in order to emphasize parallel movements. Polarity was approved by the systematic deflection at the end of utterance cf. e.g. [16].

As a general Gx pattern, we consistently observed a local maximum around the time of the prevoicing onset. Another maximum (which was often somewhat lower) occurred at the time of the release. This was rapidly followed by a minimum after the vowel onset. This pattern can be explained by an upward movement of the entire anatomical chain (larynx–hyoid–jaw–tongue-root–velum) during the oral (stop) closure and a following drop after the release and vowel onset. The triangle described by prevoicing onset, release point and negative curve elongation in between was then targeted to trace active lowering of the larynx, as suspected in /b/, /g/ and /g͡b/ (Fig. 3) in three subjects (M3, F1, F2). The areas were larger for /g͡b/ and /k͡p/ compared to /b/, /g/ and /k/

(p=0.040), whereas the area for /g͡b/ tends to be largest but not significantly different from that for /k͡p/. Moreover, for one speaker (M3) the area size correlated (negatively) with the intensity of (the last 5th of) prevoicing (r=-0.38, p=0.0002) but not with the duration of prevoicing, suggesting that the vertical larynx displacement in fact acts as a voicing enhancement mechanism.

Figure 3: EGG Gx patterns of averaged signals (left) for five stops (rows) of 1 female subject (columns); ±200ms from release point (vertical center line); Area size of triangle area (right).



4. DISCUSSION

Even though it has been said that /g͡b/ and /k͡p/ are related to each other the same way that English /b/ and /p/ are related to each other e.g. [17] p. 5, we have found /k͡p/ to be realized almost exclusively with prevoicing, albeit shorter than with the other voiced stops. Despite this prevoicing, /k͡p/ patterns with /k/ in certain respects: it leads to an increase in the f0 of the following vowel and it leads to vowel shortening.

Moreover, we found that voiced stops in Yoruba seem to have a voicing enhancement mechanism as suggested by the EGG Gx patterns, and in particular by the EGG / voicing intensity correlations. However, further tests need to be done in order to validate our inference from EGG Gx patterns to larynx displacement. In order to act as a voicing enhancement mechanism, the larynx would have to move down (increasing cavity size), and in this case, we hope to be able to time indicators of this movement with indicators of implosion from the airflow measurements. Crucially, the fact that we have these other measurements available to us means that the hypothesis that there is a voicing enhancement mechanism is testable.

5. ACKNOWLEDGEMENTS

We thank Bernard Comrie for his support. We are grateful to Roger Mundry, Daniel Voigt, Mario Etzrodt, Leo Lancia and an anonymous reviewer for helpful comments and suggestions. Special thanks to our informants (Joseph, Oni, Mary, James, and Florence).

6. REFERENCES

[1] Awobuluyi, Ọ. 1979. *Essentials of Yoruba Grammar*. Ibadan: University Press.

[2] Baayen, R.H. 2009. *LanguageR: Data Sets and Functions with "Analyzing Linguistic Data: A Practical Introduction to Statistics"*. R package version 0.955.

[3] Bates, D.M., Maechler, M. 2009. *Lme4: Linear Mixed-Effects Models Using S4 Classes*. R package version 0.999375-32.

[4] Boersma, P., Weenink, D. 2011. Praat: Doing phonetics by computer. Version 5.2.12.

[5] Cahill, M. 1999. Aspects of the phonology of labial-velar stops. *Studies in African Languages* 28, 155-184.

[6] Cahill, M. 2006. Perception of Yoruba word-initial [g͡b] and [b]. In Arasanyin, O.F., Pemberton, M.A. (eds.), *Selected Proc. of the 36th Annual Conference on African Linguistics*, Somerville, MA: Cascadilla, 37-41.

[7] Cahill, M. 2008. Why labial-velar stops merge to /gb/. *Phonology* 25, 379-398.

[8] Connell, B. 1994. The structure of labial-velar stops. *Journal of Phonetics* 22, 441-476.

[9] Demolin, D. 1991. Les consonnes labio-vélaire du Mangbétu. *Pholia* 6, 85-105.

[10] Demolin, D. 1995. The phonetics and phonology of glottalized consonants in Lendu. In Connell, A. (ed.), *Phonology and Phonetic Evidence, Papers in Laboratory Phonology* Cambridge Univ Press, 4, 368-385.

[11] Dogil, G. 1988. On the acoustic structure of multiply articulated stop consonants (labio-velars). *Wiener Linguistische Gazette* 42-43, 3-55.

[12] Ladefoged, P. 1964. *A Phonetic Study of West African Languages: An Auditory-Instrumental Survey*. Cambridge: Cambridge University Press.

[13] Maddieson, I. 1993. Investigating Ewe articulations with electromagnetic articulography. *UCLA WPP* 72, 116-138.

[14] Ogunbọwale, P.O. 1970. *The Essentials of the Yoruba Language*. London: University of London Press.

[15] Ohala, J.J. 1983. The origin of sound patterns in vocal tract constraints. In MacNeilage, P.F. (ed.), *The Production of Speech*. New York: Springer, 189-216.

[16] Pabst, F., Sundberg, J. 1992. Tracking multi-channel electroglottograph measurement of larynx height in singers. *STL-QPSR*, 33(2-3), 67-78.

[17] Rowlands, E.C. 1969. *Yoruba*. New York: David McKay Company.