18

Chapter 1

Labial-Velar Stop Place Identification in Igbo

Roslyn Burns
Yale University

Jason Shaw
Yale University

Previous research shows that simplex dorsals in the context of /i/-like vowels can be misperceived as coronal, a pattern which is more likely for voiceless segments than for voiced ones. We explore the effect of vowel context on the perception of labial-velar stops, pursing the hypothesis that the pattern observed for simplex velar stops will generalize to complex labial-velars. We opted for a cross-language perception study with Igbo participants listening to stimuli produced by a Nupe speaker. Our results show that while listeners mostly mispercieved labial-velars as labials, voiced labial-velars are sometimes perceived as coronal before /i/-like vowels, similar to what is observed for simplex dorsals. Voiceless labial-velars, on the other hand, did not exhibit this pattern.

1 Introduction

Within West Africa, labial-velar plosives /g͡b/ and /k͡p/ often exhibit distributional restrictions based on vowel context. For example, in Dagbani (Niger-Congo: Gur), labial-velars are reported to become labial-palatals in the context before front vowels as shown in example (1) with [i] and [a].

Change author. Change title. To appear in: Change Volume Editor. Change volume title. Berlin: Language Science Press. [preliminary page numbering]

(1)	Labial-velar		Labial-palatal		(Hudu 2014, Bennett 2014)
	k͡páŋ́	‘guinea fowl’	c͡pín–î	‘guinea fowls.’	
	(< k͡páŋ–gá)				
	g͡báŋ	‘squat’	ɟ͡bíŋm–g͡báŋ	‘elephant grass’	

In (1), words sharing the same phonological form in the root undergo a vowel alternation to [i] in specific morphological contexts. When the vowel raises, it triggers palatalization on the preceding labial-velar resulting in a labial-palatal segment. The pattern of labial-velars becoming labial-palatal applies more broadly to all front vowels in Dagbani, not only [i] (Bennett 2014: 116).

In Fang (Niger-Congo: Bantoid), however, the high front vowel /i/ reportedly triggers labial-velars in noun class prefixes to become a simplex labial as shown in example (2).

(2)	Labial-velar		Simplex Labial		(Mve et al. 2019, Burns 2023)
	k͡p–ú	‘my cl.2’	p–í	‘his/her cl.2’	
	k͡p–ɛ́	‘your sg. cl.2’			
	k͡p–əsə́	‘our cl.2’			
	k͡p–ənə́	‘your pl. cl.2’			

In (2), the class 2 prefix /k͡p–/ before vowel-initial possessive roots has two forms: [k͡p–] and [p–]. The [p–] allomorph only occurs before the high front vowel [i] which is known to be the most aggressive trigger of palatalization, yet instead of finding a labio-palatal in this environment, like in Dagbani, we find a simplex labial. Burns (2023) argues that an additional allomorph of /k͡p/ is found before consonant-initial roots, [bə–] < * k͡pi–. The reflex of this allmorph also has a simplex labial which developed from a labial-velar before a high front vowel.[1]

In addition to distributional restrictions and phonological alternations, like those shown above, upon conducting a preliminary survey of two major Niger-Congo languages spoken in Nigeria, we observed under-attestation of labial-velars in some vowel contexts. The static distributions of labial-velars across the lexicon suggest that labial-velars are less common in similar vowel contexts across Yoruba (Niger-Congo: Yoruboid, University Press 2001) and Igbo (Niger-Congo: Igboid, Echeruo 1998, Igwe 1999). Specifically, in Yoruba labial-velars were underattested before both /i/ and /u/ (University Press 2001) whereas in Igbo they are under-attested before high front vowels (Echeruo 1998, Igwe 1999).

---

[1] In Fang, class prefix vowels neutralize to [ə] (Burns 2023: 22).

Production and perception biases often underlie the development of synchronic sound system patterns. Over time, these biases may lead to the diachronic development of specific distributions. Distributional patterns can be *phonologized*, resulting in a grammatical rule that generalizes to new forms. In this way, production and perception biases can serve as the historical source of phonological alternations of the sorts found in (1) and (2) or systematic gaps in the lexicon.

In this paper, we investigate whether a perceptual bias may play a role in the development of distributional restrictions on labial-velars across different vowel environments by means of an open-response cross-language listening task. The motivation for running an experiment testing the influence of vowel context on labial-velar stop-place identification can be found in past work, both experimental and typological, on *velar palatalization*. There are many definitions of velar palatalization. For our purposes, we define it as a contextual change whereby a velar comes to be pronounced further forward in the oral cavity, as a coronal.

Perceptually, velars are confusable with coronals in certain vowel contexts. Guion (1998) found that English listeners are more likely to confuse simplex velars with coronals in the context of /i/ than in the contexts of /ɑ/ and /u/. Misperception is even more likely if the consonant is voiceless (i.e. /k/) rather than voiced (i.e. /g/). Wilson (2006) reports an artifical grammar experiment, in which participants learned velar palatalization. They found that participants exposed to velar palatalization conditioned by lower front vowels generalized the process to higher front vowels. In contrast, those exposed to velar palatalization in the environemnt of the high front vowel did not generalize the process to lower front vowels. These results were interpreted as a substantive bias towards high front vowels as palatalization triggers.

The environments in which velars are likely to be misperceived as coronal correspond to the known typological patterns of palatalization with a primary place change (Bhat 1978, Bateman 2007). (3) summarizes the typological behavior of palatalization in simplex segments, as reported in Bateman (2007).

(3) a. Triggers
    i. Lower front vowels do not trigger palatalization unless higher front vowels do.
    ii. High back vowels do not trigger palatalization unless high front vowels do.

b. Targets
   i. Voiced consonants do not undergo a primary place change unless voiceless consonants do.
   ii. Labials do not undergo a primary place change unless either coronals or dorsals do.

Notably missing from the typological behavior of targets in (3b) is what complex segments like labial-velars do. Labial-velars involve the coordination of at least two independent articulators (Ladefoged 1968): the lips and the tongue dorsum.[2] Burns (2023) hypothesizes that vowel-conditioned alternations for labial-velars have their basis in how labial-velars are misperceived in palatalizing contexts. Based on Connell's (1994) analysis of Ibibio (Niger-Congo: Cross-River) consonant place acoustics, Burns (2023) proposes that in palatalizing contexts, labial-velars may exhibit perceptual confusion with both labials and coronals. This contrasts with other vowel contexts, in which labial-velars are confused only with labials. This hypothesized pattern of confusion comes from the assumption that the tongue body will be fronted in palatalizing contexts, which has the effect of raising F2 in the formant transition, a typical cue to consonant place.

There are relatively few perceptual studies on West African languages. Some recent examples are Rose et al. (2023) and Ozburn et al. (to appear) on ATR harmony. We know of only one study on labial-velar perception, which reported data from just a handful of Yoruba speakers (Cahill 2006). Given the relative paucity of perceptual studies, our starting point in investigating labial-velars in different vowel contexts is centered around the typological literature on palatalization. To this end, our perceptual survey assess the questions in (4).

(4) a. Do labial-velars exhibit the most errors before /i/ (and progressively fewer as one moves away from the ideal palatalization trigger)?
   b. Do voiceless labial-velars exhibit more errors than voiced labial-velars?
   c. Do listeners believe that they are hearing a coronal in these contexts?
   d. Do the patterns of confusion for simplex components of labial-velars mirror the behavior of labial-velars?

If the answers to the research questions (4a-c) are affirmative and the same confusion patterns are observed with simplex velars, but not simplex labials, then it

---

[2] Labial-velars fall into three major airstream mechanism types, two of which also involve coordinated laryngeal gestures.

is likely that the synchronic patterns described above, e.g. (1) and (2), are due in part to the misperception of the dorsal constriction of the labial-velar.

The rest of this paper is organized as follows. §2.1 provides background information on the two languages involved in the cross-language study, Nupe and Igbo, focusing specifically on the differences in the consonant systems. §2.2 describes the Nupe materials used as stimuli. §2.3 describes the procedure that Igbo participants followed and our coding of the responses. §2.4 reports our methods of statistical analysis. §3 presents the findings of our study and how they compare to the questions presented in (4) above. §3.1 presents how vowel place influences the consonant place responses for labial-velars, thereby addressing (4a-c). §3.2 presents how vowel place influences consonant place responses for simplex labials and simplex velars and how these patterns compare to the labial-velars patterns, thereby addressing (4d). §3.3 summarizes how the Igbo listeners' responses for both simplex (labial and velar) and complex (labial-velar) segments align with the general palatalization typology outlined in (3) above. §4 discusses the implications of the findings for previous studies and future research. Finally, §5 closes by summarizing the overall findings of the study.

2 Methods

This section presents the methods for our cross-language study involving Nupe (Niger-Congo: Nupoid) stimuli and Igbo (Niger-Congo: Igboid) listeners. We opted for a cross-language design in order to increase the level of difficulty of the task. Although labial-velars are common across West African languages, the specific phonetic realization of these segments varies substantially across languages (Ladefoged 1968). We chose a somewhat less common language, Nupe, to be the language of the stimulus items to increase the level of difficulty of the task for Igbo listeners. Since we are interested in patterns of misperception, we needed to make the task difficult enough to induce errors. The difference between speaker language and listener language served this purpose.

2.1 Language Background

The language of the stimuli, Nupe, has approximately 800,000 speakers (Lewis et al. 2019); the listener language, Igbo has approximately 29,000,000 speakers (Lewis et al. 2019). Both Nupe and Igbo have a full set of voicing contrast across labial-velar, velar, and labial places: i.e. /k͡p, g͡b, k, g, p, b/. Nupe has a 5-vowel system and Igbo has an 8-vowel system with additional nasalized vowel contrasts (Ikekeonwu 1991, Uguru 2015, Moran & McCloy 2019).[3] In some varieties of Igbo, the labial-velars are reportedly plosives; in others, they are often realized as implosives /ɓ̥, ɓ/ (Moran & McCloy 2019, Ukaegbu 2023). The development of bilabial implosives from labial-velars is a common innovation in West Africa (Ladefoged 1968, Ponelis 1974) as the double articulation is sometimes coupled with a larynx lowering gesture and incomplete dorsal closure.

We make the assumption that Igbo listeners will hear the Nupe labial-velar as either a labial-velar with the same voicing or as a bilabial implosive with the same voicing.[4]

2.2 Materials

We developed a list of 96 disyllabic nonce words of the shape CVCV. These words were read by Nupe native-speaker linguist Ahmadu Kawu. The first C was either a labial, coronal, dorsal, or labial-velar plosive and varied in voicing (either voiced or voiceless). The first vowel was either /i, e, a, u/ and varied in tone (either low or high). The second syllable always began with a coronal and always ended in the vowel /a/ matching in tone with the first syllable. Kawu read these words in two Nupe carrier phrases outlined in (5).

(5) a. Wun ganan ___ be.
3sg speak ___ again.
'He said ___ again.'
b. Wun ka ___ be.
3sg write ___ again.
'He wrote ___ again.'

Kawu recorded the phrases on a Zoom H4n digital recorder at a sample rate of 44.1 khz.

Recordings were annotated in Praat (Boersma & Weenink 2019) and aligned using the Montreal Forced Aligner (McAuliffe et al. 2017) with the English pro-

---

[3] In addition to these plosive contrasts, Igbo has aspiration contrasts, palatalization contrasts in labials, and labial-velarization contrasts in velars.

[4] While voiceless implosives exhibit prevoicing, they are still differentiated from voiced implosives in having a 20–50 ms period of silence prior to the prevoicing (McLaughlin 2005: 207–209). As discussed in §2.2, based on how the stimuli for the experiment were processed, prevoicing cannot be a cue that listeners use to interpret the behavior of plosives in this study.

nunciation model and Nupe dictionary created by the researchers. All forced alignment was manually checked and segment boundaris were moved to the nearest zero-crossing. Of the 192 target utterances, we selected 96 tokens for inclusion in the experiment based on (a) vowel peripherality (avoiding heavily centralized utterances), (b) duration (avoiding abnormally long or short tokens), (c) intensity (avoiding abnormally soft tokens), and (d) clarity.[5] After selecting the tokens for inclusion, we extracted the initial CV of the nonce word from the stop burst to the offset of the vowel. This method removed prevoicing and stop closures from the stimuli. Kawu produced a labial-velar glide instead of a plosive in the token /gu/, i.e. [wu] instead of [gu]. For this item, we excised the portion from the onset of glide voicing to the offset of the vowel. After extracting the initial CV, which was always less than 140 ms in duration, amplitude was normalized across all tokens in Praat to 70 dB.[6]

2.3 Procedures

51 Igbo listeners were recruited to participate in an experiment hosted on the Gorilla platform (https://gorilla.sc/). All consent forms and prompts were written in Igbo, translated from English to Igbo by Dr. Carol Anyagwa of the University of Lagos. One participant was excluded for failure to complete the task. After consenting, participants were instructed to listen to a steady tone (75 Hz, 70 dB) and adjust their speakers to a comfortable volume. After this, they were instructed to listen to the normalized CV stimuli (described in §2.2) and type a response indicating what they heard. Participants completed 7 practice trials with stimuli not used in the experiment, and then completed 96 randomized experimental trials. After completing the task, they were asked to fill out a demographic survey.

Responses were downloaded in a tab-delineated file and checked for interpretability (n = 4800). Answers with no response (n = 111), commentary about not being able to hear (n = 3), and uninterpretable orthographic representations (e.g. <q>, <c>, etc., n = 18) were excluded from analysis. In total, 4668 usable tokens were coded for voicing and response place based on the orthographic representation. The response places were coded as labial, coronal, dorsal, labial-velar, or

---

[5] Linear models evaluating the effect of carrier phrase on formant frequencies returned no significant effect of carrier phrase.

[6] We performed acoustic analysis for stop place cues including investigating the maximum frequency at maximum amplitude in burst, F2 at 7% of the CV duration, and the change in F2 from 7% to 14% of the CV duration. For the results of this analysis see Burns & Shaw (2023).

neither. (6) shows the break-down of how each orthographic symbol was coded for place.

(6) Labial: f, v, m, b, p
Coronal: t, d, ch, l, r, y
Dorsal: k, g
Labial-velar: w, kp, gb, kw, gw, vb[7]
Neither: Vowel-initial, h-initial

Errors were coded based on whether the place represented in the response matched the Nupe stimuli place features. Responses to /gu/ were coded as correct if Igbo listeners wrote either a velar or a labial-velar as Kawu produced [w] instead of [g] in these contexts (see §2.2).

2.4 Analysis

In order to evaluate the statistical reliability of the trends in the data, we fitted logistic regression models with mixed effects. We first investigated error rate. The dependent variable in this case was consonant place identification accuracy, coded as 0 for incorrect and 1 for correct. Our models contained a random intercept for participant. Our fixed effects coded for properties of the stimulus items: consonant place labial, velar, labial-velar, consonant voicing voiceless, voiced, and vowel /a/, /e/, /i/, /u/. We also tested for two-way interactions between consonant place * vowel and between consonant voicing * vowel and for the three-way interaction between consonant place * consonant voicing * vowel.

We evaluated statistical significance through nested model comparison, via ANOVA, and consideration of the Akaike Information Criterion (AIC). We considered a fixed factor to be statistically reliable if it both explained significantly more variance than a model that lacked it and lowered the AIC. In the case of significant interactions, we ran post-hoc models on each level of our fixed factors. For example, separate models with vowel as a fixed factor were fit to each combination of place and voicing, e.g. a model for /g͡b/, /k͡p/; /g/, /k/; /b/, /p/; and /d/, /t/. All of our fixed factors were treatment coded. To assess the effect of each level of a fixed factor, such as vowel, we re-ordered the levels so that a different

---

[7] We treated <vb> as a typographical error for <gb> given the proximity of the <g> and <v> keys.

level served as the intercept. The level of a fixed factor was taken to be statistically reliable if it had a p-value equal to or below 0.05.[8]

In addition to error rate, we also fit logistic models assessing the probability of a coronal response, with coronal coded for 1 and other responses coded for 0. These models followed the same structure and procedure as those fit to error rate (described above).

3 Findings

This section presents the results of our experiment. The presentation follows the research questions in (4). We first examine research questions (4a-c) for labial-velars (§3.1) and then (4d) for the simplex components of the labial-velars (§3.2). We close the section with a summary of the findings (§3.3).

3.1 Labial-Velar Error Rates and Consonant Place Responses

Among responses to labial-velar stimuli (/g͡b/ = 1165, /k͡p/ = 1171), /g͡b/ showed fewer errors than /k͡p/ (/g͡b/ errors = 372, /k͡p/ errors = 739). Figure 1 shows the error rates for voiced labial-velars (left panel) and voiceless labial-velars (right panel) in each vowel context. Error bars represent the standard error calculated across speakers. Statistical significance, determined according to the procedure described in §2.4, is indicated with an asterisk.

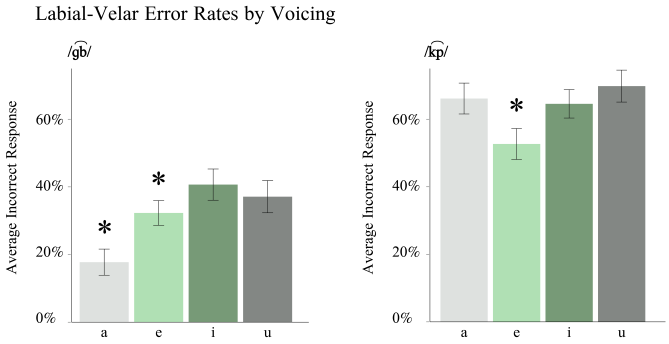

Figure 1: Average error rates for labial-velar stimuli

---

[8] P-values are based on an F-test using the Satterwaithe approximation method in the LMERTEST package.

As shown in the left panel of Figure 1, for /g͡b/, /i/ has the highest error rate. Among the front vowels, there is a significant error rate decreases from /i/ to /e/ and then from /e/ to /a/ as indicated by the asterisk. Thus, among the front vowels, there is a gradual decrease in error rates with /i/ exhibiting the most errors and fewer errors as the front vowel becomes lower. Within the high vowels, /i/ has a numerically higher error rate than /u/ but this difference was not statistically reliable.

Among responses for /k͡p/, shown in the right panel of Figure 1, the difference in error rate between /i/, /a/, and /u/ was not statistically reliable. /e/, however, had a reliably lower error rate compared to all other vowels. In this respect, only /g͡b/ exhibits error patterns consistent with the previous literature on palatalization.

Turning next to the outcomes of erroneous responses, we find that both voiced and voiceless labial-velars are most frequently confused with simplex labials. Figure 2 presents the percentage of error types by voicing (separate panels) and vowel context (x-axis).

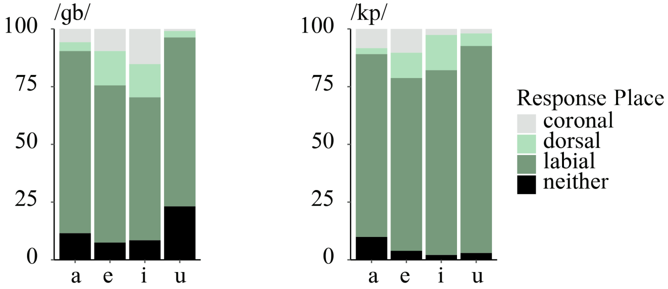

Figure 2: Proportions of incorrect labial-velar response types

In Figure 2, we can see that simplex labial responses comprise over 60% of the errors in both /g͡b/ (n = 257) and /k͡p/ (n = 601) across all vowel contexts. For /g͡b/, the /i/context triggers a higher percentage of coronal errors (n = 31) than either /e/ or /a/, a difference which was statistically significant. There is also a difference in coronal responses between /e/ and /a/, although this difference was not statistically reliable in our sample. Similarly, coronal responses for /i/ make up a higher percentage of incorrect responses than coronal responses for /u/ (also statistically significant). For /k͡p/ coronal errors (n = 41), /i/ and /u/ pattern together (i.e. are not statistically different) with the lowest incidence of coronal errors. The other two vowel contexts, /a/ and /e/, also pattern together (not significantly different from each other), showing significantly more coronal errors than /i/ and /u/. This suggests that while coronal errors for /g͡b/ follow the typological patterns described for palatalization (i.e. lower front vowels do not trigger palatalization unless higher front vowels do, high back vowels do not trigger palatalization unless high front vowels do), /k͡p/ errors do not.

To follow up the differential effect of vowel context on voiced vs. voiceless labial-velars, we looked at the error types based on the listener's perception of voicing. This perspective on the data revealed a striking pattern. Figure 3 below shows the place errors of labial-velars grouped by the stimulus voicing property (columns) and the listener's accuracy in voicing perception (rows).[9]

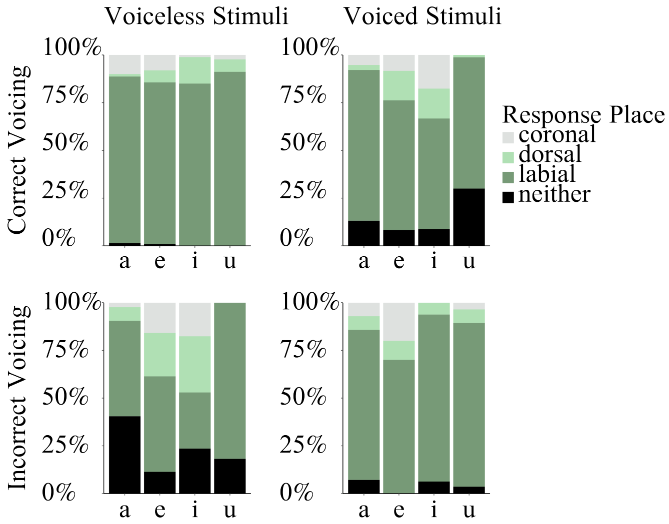

Figure 3: Proportions of incorrect labial-velar response types based on voicing perception

As shown in the first row, if listeners hear /g͡b/ and correctly perceive that the consonant is voiced (n = 304), the pattern observed is the one wherein /i/ has the highest percentage of coronal responses and there are progressively fewer as the vowel moves away from /i/ (top right panel). If listeners hear /k͡p/ and correctly perceive that the consonant is voiceless (n = 603), /i/ exhibits no preference for coronal responses (top left panel). Looking at the second row, however, if listeners hear /k͡p/ and perceive it as voiced (i.e. /g͡b/, n = 136), the pattern

---

[9] These were not tested for statistical significance due to low occurrences across the different categories.

emerges wherein /i/ receives the highest percentage of coronal responses and other vowels exhibit progressively fewer as they move away from this ideal trigger of palatalization (bottom right panel). When listeners hear /g͡b/ and perceive it as voiceless (i.e. /k͡p/, n = 68) there is no preference for coronal responses before /i/ (bottom left panel). These findings suggest that the perception of voicing is important to how Igbo listeners process labial-velars in palatalizing contexts. Interestingly, the effect of perceived voicing patterns differently than expected from past work on the role of voicing in velar palatalization. We return to this result in the discussion (§3.3).

3.2 Simplex Consonant Errors

We close the presentation of the results with errors among simplex labial and simplex velar stimuli. Figure 4 shows the error rate for voiced (left panel) and voiceless (right panel) simplex stimuli by vowel context. The top row shows velar stimuli and the bottom row shows labial stimuli.

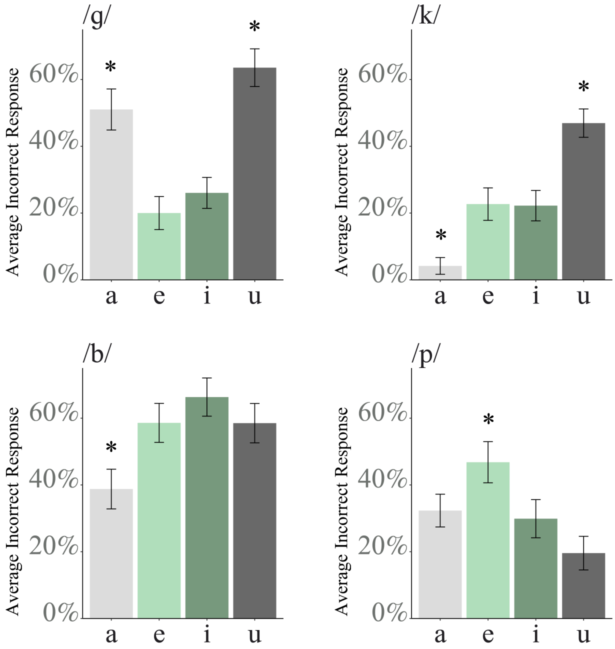

Figure 4: Simplex dorsal and simplex labial error counts

Statistically, for simplex /g/ (responses = 390, errors = 156), /i/ and /e/ pattern together with the fewest overall errors. /a/ had significantly more errors than other front vowels, /e/ and /i/. /u/ had significantly more errors than /a/, the most errors overall. For simplex /k/ (responses = 390, errors = 94), /i/ and /e/ both have more errors than /a/, a difference which is statistically significant. /u/ has significantly more errors than the other contexts. For simplex /b/ (responses = 386, errors = 214), /a/ has the lowest error rate while /i/, /e/, and /u/ pattern together with a higher error rate. For simplex /p/ (responses = 387, errors = 124), /e/ has the highest error rate, a difference which was statistically significant. /i/ patterns together with /a/ and /u/ with lower error rates. In sum, of the four simplex stops examined, only voiced labials showed high error rates for /i/, which is typologically the most aggressive trigger of palatalization. The voiceless labial has the high error rates for /e/, but no other vowel. Among velars, either /a/ or /u/ exhibit the highest error rates.

In order to interpret how the error rate patterns relate to the palatalization hierarchy described in (3), we need to know in particular how coronal responses are distributed across the different place, vowel, and voicing combinations. Figure 5 shows the rate of different place responses among the errors for simplex velars and simplex labials.

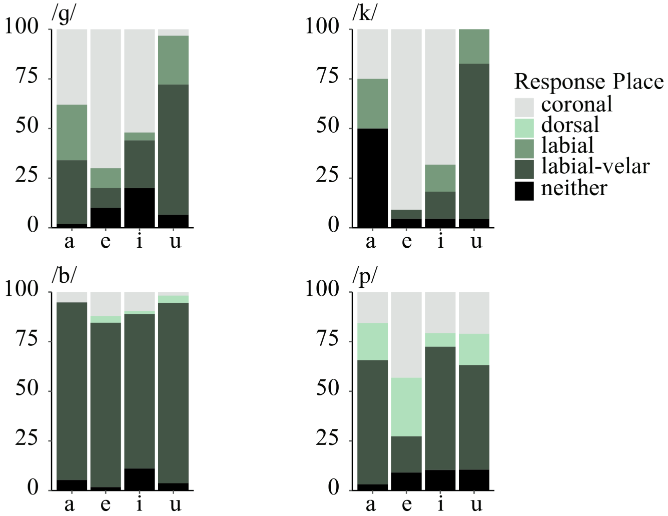

Figure 5: Proportions of incorrect simplex dorsal and simplex labial response types

For simplex /g/ coronal errors (n = 49), all front vowels pattern together statistically having higher rates of coronal responses than /u/. For simplex /k/ coronal errors (n = 36), /i/ and /e/ statistically exhibit the most coronal responses. There are significantly fewer coronal responses before /a/ and /u/ . In this respect, both /g/ and /k/ follow the palatalization hierarchy, but in different ways. For both, lower front vowels do not condition the percept of coronals to the exclusion of higher front vowels. This is consistent with the palatalization hierarchy. Additionally, the high back vowel does not condition the percept of coronals to the exclusion of the high front vowel, which is also consistent with the palatalization hierarchy. The difference between the two simplex velars is that coronal responses are more balanced across the front vowels for the voiced velar /g/ than for the voiceless velar /k/, which we did not expect.

When we look at the error pattern for simplex labials, the other closure component of the labial-velar, we find that with simplex /b/ coronal errors (n = 16), there is statistically no preferred vowel environment. For simplex /p/ coronal errors (n = 34), there were statistically more coronal responses in the context of /e/ than in the other vowel contexts. This indicates that simplex labials do not follow the patterns that have been described for the palatalization hierarchy outlined in (3), but simplex velars do.

3.3 Summary

The previous typological literature on palatalization identified target place, target voicing, and trigger vowel as important factors in observing palatalization (see 3). Our investigation of how complex dorsals (i.e. labial-velars) factor into the known typology indicates that while vowel context is important to how listeners perceive labial-velars, there is a split in target voicing properties which runs counter to the known typological behavior. That is, the voiced labial-velar /g͡b/ is more consistent with the palatalization hierarchy than /k͡p/. This behavior appears to be connected specifically to how listeners perceive voicing, as opposed to the actual voicing of the segment.

When comparing complex dorsals (i.e. /g͡b/ and /k͡p/) to their simplex closure components (i.e. /g/, /k/, /b/, /p/), we find that /g͡b/, /g/, and /k/ statistically exhibit behaviors consistent with the palatalization hierarchy concerning coronal response rates. Neither /k͡p/, /b/, nor /p/ exhibit error patterns consistent with the literature on palatalization triggers. While /b/ exhibited no preferences by vowel for coronal responses, both /k͡p/ and /p/ showed a preference for coronal responses before /e/ but never /i/ (with /k͡p/, /e/ and /a/ patterned together).

4 Discussion

We sought to understand if misperception of labial-velars in certain vowel contexts mirrored patterns of underattestation in those same contexts. We tested this hypothesis through the lens of velar palatalization. Underattestation of labial-velars occurred in the same environments which have been reported to trigger changes from simplex dorsals to coronals, in both experimental and typological literature.

The results of our cross-language Nupe-Igbo study found that all labial-velars are associated with high confusability with simplex labials regardless of vowel context. When looking specifically at confusions of labial-velars as coronal consonants, we found patterns reminiscent of velar palatalization. These findings provide some support for Burns' (2023) hypothesis that labial-velars may have two outcomes in palatalizing contexts. In addition to becoming coronal in these contexts, they may also simplify to simplex labials. The latter outcome may develop to avoid confusing lexical items with labial-velars for lexical items for coronals, as labial-velar confusability with coronals increases specifically in palatalizing contexts. A change from a labial-velar to a coronal would increase contrast with a simplex labial in a palatalization context. As simplex labials are already confusable with labial-velars independent of vowel context (and vice versa), there is not much that is lost from an identification standpoint when a labial-velar is replaced by a simplex labial, but there is an added benefit in that simplex labials are not notably more confusable with coronals in palatalizing contexts. Thus, the alternation of labial-velars with labials in palatalizing contexts may support system-wide maintenance of contrast.

Although our study found that vowel contexts known to condition velar palatalization also conditioned the percept of a labial-velar consonant as coronal, this pattern surfaced for the voiced labial-velar but not the voiceless labial-velar, in an apparent reversal of the voicing hierarchy observed among simplex velars in previous literature. While Igbo perception of simplex velars did exhibit responses that are consistent with the palatalization hierarchy, simplex /g/ appears to exhibit more environments of palatalization (i.e. all front vowels) than simplex /k/. This follows the same apparent voicing reversal as observed for labial-velars. It should be noted that neither simplex labial conformed to the known behaviors of the typological literature on velar palatalization. This is in line with the target place hierarchy, wherein either coronals or dorsal may undergo the process of palatalization to the exclusion of labials, but the inverse does not occur. Interestingly, /k͡p/ appeared to pattern most closely with /p/. The split in the behavior of complex segments, where /g͡b/ patterns with velars and /k͡p/ patterns with labials, leads to the question of whether or not there is some underlying motivation that would lead us to expect complex classes to exhibit disunity in which simplex component of the complex gesture that they pattern with.

Another key finding from our study, regarding the voicing hierarchy reversal, is that it seems to depend on listener perception of voicing. When Igbo listeners mispreceived a Nupe voiced labial-velar stimuli as voiceless, the patterns of consonant place misidentification across vowel contexts matched the patterns found for voiceless labial-velar stimuli (3). The same was true for Nupe voiceless labial-velar stimuli voicing misperception. Voiceless labial-velar stimuli misperceived as voiced showed the patterns of consonant place misidentification across vowel contexts found for voiced labial-velar stimuli. This pattern could possibly relate to the distribution of voiced vs. voiceless labial-velars across vowel contexts in the lexicon. A cursory examination of these patterns, based on counts in two Igbo sources (Igwe 1999, Echeruo 1998), showed that there were differences by voicing. Voiced labial-velars occurred less often before /i/ than voiceless labial-velars. The same was true for the /e/ environment; voiced labial-velars occur less often before /e/ than voiceless labial-velars. Possibly, these distributional facts influenced listener expectations in our study. Listeners that perceived the vowel /i/ or /e/ and voicing may be biased against responding labial-velar, since voiced labial-velars are under-attested in this environment. [10]

One possible direction for future research is consideration of variation in laryngeal properties found across Nupe and Igbo. Many varieties of Igbo have a glottal gesture that is coupled with the oral gestures (see the discussion of phonemic system alignment in §2.1). In some varieties, the pairing of these articulatory gestures is such that /g͡b/ > /ɓ/ and /kp/ > /ɓ̥/. The primary cue differentiating voiceless implosives from voiced implosives is a short period of silence during the closure (Ladefoged 1976 as cited in McLaughlin 2005). This is different from the Nupe cues to voicing in labial-velars, which lack voicing during closure, at least in our materials, and contain robust cues to voicing in the burst. The perceptual stimuli we used in this study lack closure information that would allow Igbo listeners to identify the difference between voiced and voiceless implosives using the primary cue of voicing duration during the closure. Our stimuli did, however, retain burst information. For Igbo listeners with the voiced and voiceless implosive distinction, the burst information may differentiate implosives (both

---

[10] An anonymous reviewer noted that many West African languages lack labial-velars before rounded vowels likely due to their development from *KUV > K͡PV.

voiced and voiceless) from non-implosives. Thus, a cue to voicing in our materials may be reinterpreted as a cue to ingressive airstream for some Igbo users. Any reinterpretation of burst cues for voicing might interact with the effect of vowel quality on stop-place perception because of the distribution of voiced and voiceless labial-velars across vowels, as discussed above.

A possible implication of the differing cues to voicing is that voiceless /k͡p/ in Nupe could be perceived by certain Igbo listeners as /p/. On this note, there were some similarities in how the effect of vowel context patterned across /k͡p/ and /p/. For these two consonants, place identification accuracy was similar across /i/, /a/, and /u/ with /e/ being statistically different. /e/ showed higher place identification accuracy for /k͡p/ (Figure 1) and lower place identification accuracy for /p/ (Figure 4). Unfortunately, given that simplex dorsals also exhibit reversal of the voicing hierarchy, further studies, both perceptual and articulatory, are needed to determine the exact motivations underlying all voicing reversals. As a first step, it may be informative to conduct a version of our study which provides cues to voicing during closure to see if the voicing hierarchy reversals persist even with more robust cues to voicing. Past work has shown that voicing during closure is an important cue to laryngeal contrasts in some languages with labial-velars (Grawundera et al. 2011) and, moreover, excising this cue to voicing from the speech signal can lead to stop place misidentification (Cahill 2006).

5 Conclusion

We presented the results of a cross-language perception study investigating how vowel context affects stop place perception. In an open response task, Igbo listeners heard eight Nupe stops, /k͡p, g͡b, k, g, p, b/, in four different vowel environments, /a, e, i, u/. Our primary theoretical motivation was a test of whether misperception of labial-velars as coronal would follow the three implicational relationships identified for velar palatalization. Results indicated that voiced labial-velars do indeed follow the velar palatalization hierarchy. There are more coronal percepts of voiced labial-velars in the environment of /i/ followed by /e/ and then /a/. Voiceless labial-velars, on the other hand, did not show this same pattern. This was surprising because it had been claimed that, for simplex velars, voiceless stops are more likely to palatalize than voiced stops. However, simplex velars in Igbo also flouted this generalization about voicing with /g/ exhibiting more coronal errors than /k/. Additionally, the coronal errors for /g/ were more balanced across all front vowel contexts than for /k/, which was also unexpected.

A complete account of the voicing hierarchy reversals reported here requires future research, but our results indicate that voicing perception itself plays a key role in conditioning how vowel context impacts stop place perception.

References

Bateman, Nicoleta. 2007. *A crosslinguistic investigation of palatalization*. University of California San Diego. (Doctoral dissertation).

Bennett, William. 2014. Some differences between clicks and labio-velars. *South African Journal of African Languages* 34(2). 115–126.

Bhat, Dharbe N.S. 1978. A general study of palatalization. In J. Greenberg (ed.), *Universals of Human Language*. Vol. 2, 47–92. Stanford: Stanford University Press.

Boersma, Paul & David Weenink. 2019. *Praat: doing phonetics by computer*. Comp. software. Version 6.0.46. praat.org.

Burns, Roslyn. 2023. Balancing social determinism vs. sound change: the case of fang. *Journal of Historical Linguistics*. DOI: 10.1075/jhl.22012.bur.

Burns, Roslyn & Jason Shaw. 2023. Effect of vowel context on stop place identifcation in yoruba. *Proceedings of the International Conference on Phonetic Sciences* 20. 2946–2950.

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

Connell, Bruce. 1994. The structure of labial-velar stops. *Journal of Phonetics* 22(4). 441–476.

Echeruo, Michael J. C. 1998. *Igbo-english dictionary: a comprehensive dictionary of the igbo language with an english-igbo index*. New Haven: Yale University Press.

Grawundera, Sven, Bodo Winter & Joseph Atoyebi. 2011. Voicing of labiovelar stops in yoruba. *Proceedings of the International Conference on Phonetic Sciences* 17. 767–770.

Guion, Susan. 1998. The role of perception in the sound change of velar palatalization. *Phonetica* 55(1–2). 18–52.

Hudu, Fusheini. 2014. What is the phonological word in dagbani? a positional faithfulness account. *Ghana Journal of Linguistics* 3(1). 1–44.

Igwe, G. E. 1999. *Igbo-english dictionary*. Ibadan: University Press.

Ikekeonwu, Clara I. 1991. Igbo. *Journal of the International Phonetic Association* 21(2). 99–101.

Kawu, Ahmadu. 2023. Nupe native-speaker linguist.

Ladefoged, Peter. 1968. *A phonetic study of west african languages: an auditory-instrumental survey*. Cambridge: Cambridge University Press.

Ladefoged, Peter. 1976. The stops of owerri igbo. *Studies in African Linguistics* 7. 147–163.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig. 2019. *Ethnologue: languages of the world [online]*.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger. 2017. Montreal forced aligner: trainable text-speech alignment using kaldi. In *Interspeech*, vol. 2017, 498–502.

McLaughlin, Fiona. 2005. Voiceless implosives in seereer-siin. *Journal of the International Phonetic Association* 35(2). 201–214.

Moran, Steven & Daniel McCloy (eds.). 2019. *Phoible 2.0*. Jena: Max Planck Institute for the Science of Human History. https://phoible.org/.

Mve, Patrick, Nelson Tschonghongei, Pierpalo Di Carlo & Jeff Good. 2019. Cultural distinctiveness and linguistic esoterogeny: The case of the Fang language of the Lower Fungom, Cameroon. In Pius W. Akumbu & Esther P. Chie (eds.), *Engagement with Africa: Linguistic essays in honor of Ngessimo M. Mutaka*, 163–178. Köln: Rüdiger Köppe Verlag.

Ozburn, Avery, Gianna F Giovio Canavesi & Samuel Akinbo. to appear. Perception of atr in dàgááre [dàgáárɪ̀]. *Laboratory Phonology*.

Ponelis, Fritz. 1974. On the dynamics of velarization and labialization: some bantu evidence. *Studies in African Linguistics* 5(1). 27–58.

Rose, Sharon, Michael Obiri-Yeboah, Sarah Creel & Sarah C. Creel. 2023. Perception of atr contrasts by akan speakers: a case of perceptual near-merger. *Laboratory Phonology* 14(1).

Uguru, Joy Oluchi. 2015. Ika igbo. *Journal of the International Phonetic Association* 45(2). 213–219.

Ukaegbu, Nkechi. 2023. Igbo native-speaker linguist.

University Press. 2001. *A dictionary of the yoruba language*. Ibadan: University Press.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30(5). 945–982.