
Yorùbá vowel deletion involves compensatory lengthening: Evidence from phonetics

Nick Danis 

Linguistics Program, Washington University in St. Louis, USA
E-mail: nsdanis@wustl.edu

Abstract

A phonetic pilot study of Yorùbá vowel deletion shows that the vowel that remains after an adjacent vowel deletes is slightly but significantly longer than a short vowel in non-deletional contexts ($p < 0.001$). In the configuration studied here, deletion occurs in the vowel of a CV verb when occurring before a V-initial direct object ($/CV_1 + V_2/ \rightarrow [CV_2]$). However, instead of full vowel deletion as it is previously analysed (e.g. Akinlabi and Oyebade 1987, Ola Orié and Pulleyblank 2002), a compensatory lengthening analysis is proposed based on this new phonetic evidence. The experiment for this study controlled for inherent vowel duration, voicing, and manner of articulation of the surrounding consonants. These results are in line with a similar result regarding Yorùbá tone (Ajíbóyè et al. 2011) in the same syntactic (verb + direct object) configuration.

Keywords: phonology, Yorùbá, deletion, compensatory lengthening, vowel duration

1. Introduction

This work examines the phonetics of a vowel deletion process in Yorùbá (Atlantic-Congo, Nigeria). Specifically, in a pilot study, the duration of an underived short vowel is compared with the duration of a vowel that remains after deletion of an adjacent vowel (called here the “remnant vowel”) in a VV sequence.

- | | | | | | |
|-----|----------------|---------------------|------------|----------------------|---------------|
| (1) | a. $/V_1 V_2/$ | $\rightarrow [V_2]$ | $/ta+ata/$ | $\rightarrow [tata]$ | ‘sell pepper’ |
| | b. $/V_3/$ | $\rightarrow [V_3]$ | $/tata/$ | $\rightarrow [tata]$ | ‘grasshopper’ |

The duration of the remnant vowel (V_2 above) is slightly but significantly longer than the duration of the underived short vowel (V_3 above). The standard phonological account treats this process as full deletion. However, due to the difference in duration, this account must be revised. A compensatory lengthening account involves the segmental material being deleted while a mora remains, causing lengthening on the remnant vowel. An incomplete neutralisation account would involve the phonetic module spelling out the phonologically short vowel as slightly longer. Both options are discussed, and ultimately Yorùbá vowel deletion is re-analysed as compensatory lengthening.

At the time of the study, the subject had lived in the United States for the previous three months, and spoke Yorùbá daily.

Forms to be elicited were compiled with the aid of the native speaker in several sessions prior to recording, though the subject was not aware of the eventual goal of the experiment. To ease in eventual segmentation, words with the vowel [a] between voiceless obstruents, preferably stops, were sought. To control for inherent durations among different vowels, only the vowel [a] is analysed. At the time of elicitation, the target sentences were given to the speaker in groups of seven sentences per page, with a total of 102 sentences, randomised throughout. The first and last sentences on each page were recorded but ignored, to control for list intonation effects, and the subject was instructed to speak naturally, as if to a friend, based on similar techniques from Broselow, Chen and Huffman (1997). Each target sentence was repeated at least five times (non-consecutively) during the elicitation. The subject was instructed to read each sentence once, with a pause in between each one, and a longer pause or break after each page. The sentences were written in standard Yorùbá orthography, with no English present.

The full wordlist used is given in (4). Vowel durations for only a subset of these are measured; these are given later in (5) and (6).

(4) **Full wordlist**

- | | | |
|----|-------|-------------------------------|
| a. | aká | ‘cripple’ |
| b. | àkàrà | ‘bean cake’ |
| c. | àkèké | ‘scorpion’ |
| d. | apá | ‘hand’ |
| e. | àpótí | ‘stool’ |
| f. | ata | ‘pepper’ |
| g. | àtíkè | ‘make-up powder’ |
| h. | bátá | ‘shoe’ |
| i. | káso | ‘pack/fold cloth’ |
| j. | pákó | ‘wood’ |
| k. | pápá | ‘field’ |
| l. | pasè | ‘beat a rhythm with the feet’ |
| m. | pupa | ‘red’ |
| n. | tata | ‘tata’ |

Recordings were made in a sound-attenuated booth in the Phonology and Field Research Laboratory at Rutgers University. The subject wore a head-mounted AKG C420 microphone connected through a digital pre-amp, and was recorded in Goldwave at 44.1kHz. The file was saved as a WAV file and segmented in Praat.

Following Francis, Ciocca and Yu (2003) and Ladefoged (2003), vowel segmentation was determined by the first zero-crossing (where the amplitude is 0) before the first regular period of a periodic signal, and vowel offset was determined in a similar way. For vowel offset, the zero-crossing at the last stable periodic signal was marked. In some cases, there were one or two extra pulses of voicing, but these were not considered part of the vowel as the waveform had lost its shape.

4. Data and results

Each target word was elicited in the frame sentence *mo ta _ lana*, ‘I sold _ yesterday’, except where indicated below; in these latter cases, other verbs were deemed more natural by the consultant.² After excluding certain forms to control for voice and manner of articulation (see, e.g., Peterson and Lehiste 1960, and Van Santen 1992), the durations of the initial vowels (in bold) in the following nouns were measured:

(5) Remnant vowels

- a. ta **ata** ‘pepper’
- b. gba **aká** ‘cripple’
- c. ta **apá** ‘hand’
- d. ta **àkèké** ‘scorpion’
- e. ta **àkàrà** ‘bean cake’
- f. ta **àpótí** ‘stool’
- g. ta **àtíkè** ‘make-up powder’

(6) Simple vowels

- a. ta **tata** ‘grasshopper’
- b. ta **pákó** ‘wood’
- c. ge **pápá** ‘field’

The null hypothesis under the standard phonological account is that there should be no significant difference between the duration of the vowels between the two groups. However, there is indeed a significant difference between the duration of remnant vowels and simple short vowels ($t(40.458) = -4.1821$, $p < 0.001$). This is summarised in Table 1, while the means are plotted in Figure 1. The two groups were compared via the *t.test* function in R.

Table 1: Differences in durations for remnant versus short vowels

	Remnant Vowels	Short Vowels
n =	33	15
Mean =	108 ms	96 ms
StdDev =	12 ms	8 ms
$t(40.458) =$	-4.1821	
$p =$	0.0001508	

² The verb *gba* means ‘to hit’ and *ge* means ‘to cut’.

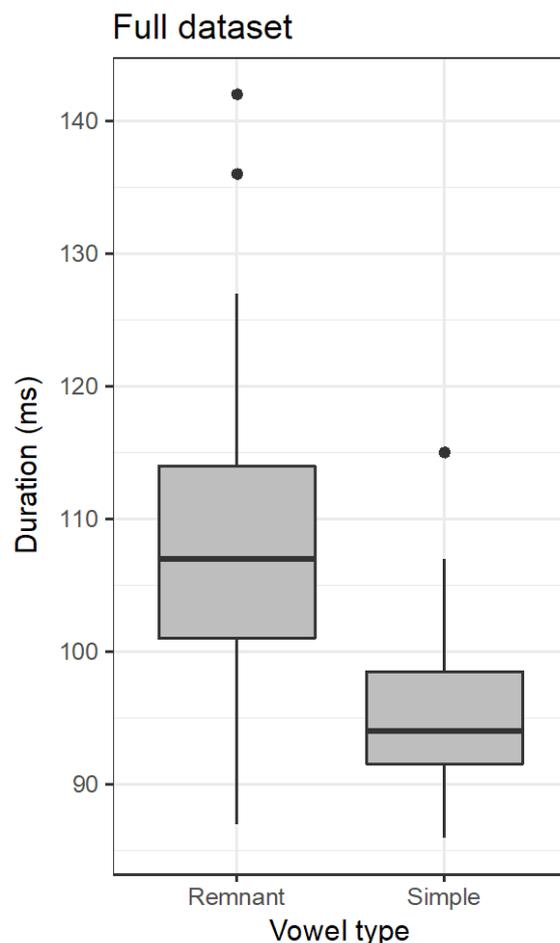


Figure 1. Means of remnant vowels versus simple vowels

On average, the mean duration of the remnant vowels is about 12 ms longer than the simple vowels. While the two sets of words are balanced for place of articulation in voicing, there are still imbalances between them, such as number of tokens, tone, and word length. However, there is one near-minimal pair between the two groups: [ata] ‘pepper’ and [tata] ‘grasshopper’. The results for these two forms are given in Table 2.

Table 2: Differences in durations for ‘pepper’ versus ‘grasshopper’

	Remnant Vowels	Short Vowels
n =	4	5
Mean =	113 ms	97 ms
StdDev =	10 ms	11 ms
t(6.84) =		2.31
p =		0.055

This difference in the minimal pair is just above significance: $t(6.84) = 2.31$, $p = 0.055$. The absolute difference in means is comparable to the full dataset. More tokens of this type are likely to yield a more robust significance. These forms are identical in segments, tone, and word length. Additionally, the tone of both words is M(id), and there is no significant difference in f_0 between the two vowels ($t(4.48) = -1.03$, $p = 0.36$); the durational differences are caused

neither by phonological tone nor fundamental frequency (see, e.g., Mamadou Y. 2017 on the effects of f_0 and duration in Yorùbá).

The difference in duration between the two vowel types is just at the Just Noticeable Difference for duration. Klatt (1976) finds that differences between 10 ms and 20 ms are perceptible, while the difference in means here is 12 ms. It is unclear if this difference is actually perceptible for a native Yorùbá speaker; a perception experiment is necessary to test this.

5. Implications for phonology and phonetics

Assuming the results are robust, there are implications for either the phonological account, its phonetic implementation, or both. If we assume the standard phonological account, where the prosodic structures containing the remnant vowel and the simple vowel are identical, then this is a potential case of *incomplete neutralisation*: two forms that should otherwise be identical have slight phonetic differences. In this case, a /VV/ triggers full deletion, resulting in a monomoraic remnant vowel V, which is realised differently by the phonetics from a structurally identical simple vowel.

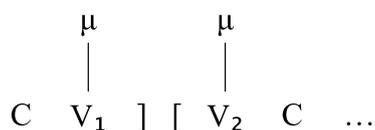
Alternatively, because of the phonetic differences found, the other option is a revised phonological account where the process is no longer treated as full deletion, but rather as root node deletion plus compensatory lengthening: the remnant vowel will lengthen due to being reassociated to the mora of the deleted vowel root node. However, this account is not completely straightforward, as the phonetic output of compensatory lengthening usually results in a vowel of which the duration is similar to that of one that is phonologically long.

5.1. Revised phonological account

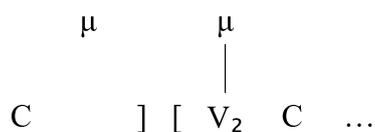
As defined by Hayes (1989), the Yorùbá facts fit the definition of compensatory lengthening, which is “[...] the lengthening of a segment triggered by the deletion or shortening of a nearby segment” (Hayes 1989: 260). Assuming the process is compensatory lengthening, the standard phonological account is thus revised. Instead of full vowel deletion, the output of the relevant phonological processes is instead root node deletion, with the mora of the underlying vowel remaining and causing lengthening on the remnant vowel:

(7) Revised phonological account

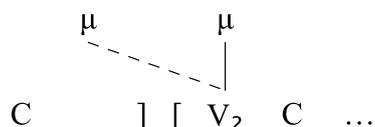
a. Underlying moraic structure



b. V_1 root node deletes



c. Remaining mora reassociates



Because the output of the phonological process is now structurally different from a simple short vowel (which projects only one mora – see the structure in (3)), there is a phonological reason why the remnant vowel should be phonetically realised longer than a simple vowel. Yorùbá has no contrastive vowel length distinctions, so the remnant vowel, even though it is only ~12 ms longer than a short vowel, is the phonetic realisation of a Yorùbá long vowel.³

This pattern is structurally identical to a more well-known case of compensatory lengthening in Luganda:

- (8) Luganda
 a. /ka+oto/ → [ko:to] ‘fireplace’ (dim.)
 b. /ka+ezi/ → [ke:zi] ‘moon’ (dim.) (Clements 1986, Goldrick 2000)

However, the crucial difference is in the phonetic realisation of the bimoraic vowel: whereas in Luganda and in other cases of compensatory lengthening the realisation is akin to a long vowel, in Yorùbá it is only slightly longer. This phonetic fact, previously unreported, is why the process in Yorùbá is usually treated as full deletion and not compensatory lengthening.

Why should the phonetic realisation of a bimoraic in Yorùbá be only slightly longer than a short vowel? De Chene and Anderson (1979) claim that “the existence of an independently-motivated length contrast in the language is a necessary condition for compensatory lengthening” (De Chene and Anderson 1979: 508). In other words, if a language does not have phonologically long vowels elsewhere, a compensatory lengthening process will not result in a phonologically long vowel. While this claim has been weakened if not rejected since (see Gess 2011 for a review), there still might be a connection between the presence of a long vowel contrast and the phonetic realisation of compensatory lengthening. Recall that Yorùbá does not have contrastive long vowels.

5.2 Phonetic implementation: Incomplete neutralisation?

It is clear from the results that the remnant vowels in Yorùbá are not as long as would be expected for bimoraic vowels cross-linguistically. However, to emphasise this point further, durations for phonologically short versus long vowels for languages with a true length distinction are given in (9).

(9) Hindi vowel durations, from Broselow, Chen and Huffman (1997)

	VV	V	VV-to-V
Speaker 1	143.5 ms	66.1 ms	2.17
Speaker 2	152.6 ms	80.4 ms	1.9

³ Thanks to Florian Lionnet for discussion on this point. Here still, further discussion of the phonetic implementation is needed to explain why the bimoraic structure is realised only slightly longer than a short vowel.

In Hindi, the long vowel is about twice as long as the short vowel, compared to an 11% difference for Yorùbá. The phonetic module would have to realise the bimoraic structure in Yorùbá well below the long vowel duration in Hindi and many other languages. However, because there are no vowel length contrasts, there is an extremely low functional load on long vowels, which might contribute to only a slight phonetic difference between short and long vowels.

An alternate approach would be to assume that the standard phonological account holds, but this is a case of incomplete neutralisation. Incomplete neutralisation describes a process where the phonological module outputs two structures assumed to be identical, but the phonetic module interprets them differently. In the present case, assuming the standard account, the phonological module outputs two phonologically short vowels, but the phonetics implements one as slightly longer.

Braver (2013: 4) defines incomplete neutralisation as instances where “the surface acoustic cues to two underlyingly distinct segments in a given context are less distinct than the segments’ canonical realizations in non-neutralizing contexts, but are not completely identical”. In the standard approach to Yorùbá vowel deletion, the simple vowels and remnant vowels are not underlyingly distinct; they are structurally the same. It is only in the revised approach that they differ, specifically in their moraic configurations. In the revised approach, it is not clear what the “canonical realisation” of a long vowel is, as there are no vowel length contrasts in the language.

Additionally, there is no clear reason where the phonetic pressure to lengthen comes from under the standard approach. Assuming a recent theory of incomplete neutralisation, Braver (2013), incomplete neutralisation occurs when there is a conflict between X, Y, and Z, where X and Y are identical phonological structures but Y and Z are words related in some paradigm.

For example, a case of incomplete neutralisation that Braver studies concerns vowel duration in Japanese, making it relevant to the facts here. A root can either occur with a particle or in isolation. When occurring with a particle, the root projects one mora; in isolation, two. However, the duration of the form in isolation is not as long phonetically as an underived bimoraic form. This is shown in (10).

(10) Japanese paradigm (Braver 2013: 127, with durations from Experiment 7)

Short	Lengthened	Underived Long
μ	$\mu\mu$	$\mu\mu$
<i>ch[i]</i> + <i>mo</i>	<i>ch[i]</i>	<i>ch[ii]</i>
‘blood’ ALSO	‘blood’	‘social.status’
54.99 ms	124.98 ms	157.45 ms

While there is only a two-way contrast in phonological structure (one mora versus two), there is a three-way contrast in phonetic duration. For lengthened form *chi*, there is a conflict between two constraints in the phonetic implementation. $DUR(\mu\mu) = TARGETDUR(\mu\mu)$ returns a lower cost of lengthened *chi* and is similar in length to underived bimoraic forms, like *chii* ‘social.status’. Additionally, there is a constraint OO-ID-DUR, which returns a lower cost the more similar the lengthened form is to the monomoraic form that occurs with the particle; this form is the base of the paradigm, as it has the highest frequency. (I refer the reader to Braver (2013) for a full and clear exposition of the data and analysis.)

What, then, does the Yorùbá paradigm look like? Still assuming the standard account, it is given in (11).

(11) Yorùbá paradigm, standard account

Short	Remnant	Underived Long
μ	μ	
t[a]ta	t+[a]ta	∅
‘grasshopper’	‘sell pepper’	
96 ms	108 ms	

Because Yorùbá has no underlying long vowels, and the standard account assumes the remnant vowel is monomoraic, there is no obvious pressure for the phonetics to realise the remnant vowel any longer than the short vowel. As one reviewer puts it, “[s]ince there is no short/long vowel length contrast in Yorùbá, a vowel of middling duration is not halfway between two contrastive categories in the language,” because the second contrastive category – long vowels – does not exist. In fact, both relevant constraints, OO-ID-DUR and (in this case) DUR(μ)=TARGETDUR(μ), return the lowest cost when the duration of the short vowel equals the duration of the remnant vowel.

The revised phonological account has a clear reason why the remnant vowel should be lengthened: it projects two moras. However, the question then becomes, why is it only lengthened about 10%? While not a case of incomplete neutralisation, it seems likely that weighted, scalar phonetic constraints of Braver (2013), Flemming (2001), and others can be used to model this process. As an anonymous reviewer points out, unlike in Japanese, the Yorùbá case does not involve competing pressures from two contrastive categories in the language, as there are no contrastive long vowels. If there is no pressure in a vowel paradigm from a contrastive long vowel, the phonetics might be compelled to realise the long vowel only slightly longer than a short vowel due to this decreased functional load. However, a full account in a gradient theory of phonetics is left to future work.

5.3 Connection to syntax

The results presented here are very similar to experimental results of two tonal processes, one in the same syntactic environment as deletion.⁴ All cases of deletion and lengthening here occur between a verb and its direct object. Ajíbóyè et al. (2011) describe a tone-lowering process in this same environment. Yorùbá has three standard tones, high (H), mid (M), and low (L), and “[in] Standard Yoruba, before a direct object [...] monosyllabic L-tone verbs are raised to M” (Ajíbóyè et al. 2011: 1634). Additionally, there is a HL contour simplification process involving the reduplicative morpheme -kí-. Examples of these processes are shown in (12) and (13), respectively.

- (12) L-raising (Ajíbóyè et al. 2011: (5d))
- a. Mo na Títí
 1 sg beat Titi
 ‘I beat titi’

⁴ Thanks to Douglas Pulleyblank for bringing this to my attention.

(13) HL-simplification (Ajíbóyè et al. 2011: (13c))

/èrò-kí-èrò/ → èròkerò
 L L H L L L L M L

What Ajíbóyè et al. (2011) find is that the M that results from HL-simplification, which they call “morphologically derived”, is identical in f₀ to an underived M tone. However, the M that results from L-raising, which they call “syntactically derived”, is significantly lower in f₀ than both the underived and morphologically derived M tone. In their analysis, the syntactic environment crucially prevents the L tone from fully deleting, and thus causes it to slightly lower the f₀ of the resulting M tone.

This makes predictions for vowel deletion as well. Vowel deletion also occurs in the morphological environments for HL-simplification – is the duration of the remaining vowel there more similar to the remnant vowels here or more similar to simple short vowels? While further experiments must be done, the explanation by Ajíbóyè et al. (2011) predicts that syntactically conditioned vowel deletion should differ from morphologically conditioned deletion.

6. Further directions

The results presented here, while significant, constitute a pilot study. More work needs to be done to conclusively show robustness across several dimensions. First, future experiments should test multiple speakers across multiple dialects.

Additionally, there are both other deletion processes in other contexts, and other types of processes in similar contexts, that should be investigated. As mentioned in Section 5.3, the vowel of a CV reduplicative particle also deletes in certain cases. Phonetically, how does the duration of the remnant vowel there relate to the remnant vowel in verb + object constructions? The interpretation of Ajíbóyè et al. (2011) suggests there will be a phonetic difference.

Between verb and vowel-initial object, there are a number of processes that can occur. The generalisation argued for in Ola Orié and Pulleyblank (2002) is that CV verbs trigger deletion, but larger verbs, such as CVCV verbs, cause vowel assimilation but not deletion.

(14) Ola Orié and Pulleyblank (2002: (3f))

jáde opó jàdo opó *jádopo ‘come out of mourning’

In the form in (14), the V₁ + V₂ sequence is resolved by changing the quality of V₁ to match V₂ instead of deleting it. While all the forms in the experimental stimulus fit the context for vowel deletion (in that the verb is CV), the quality of the deleted vowel and the remnant vowel are the same: both are [a]. This is then ambiguous between deletion and assimilation.⁵ However, the process studied here is unlikely to be *only* assimilation (resulting in a VV sequence), as Ola Orié and Pulleyblank (2002: fn. 13) state that, “at the phonetic level, sequences of identical vowels appear to be produced with a slight rearticulation. That is, in a sequence like [...e.e...], there appear to be two phonetic targets, not one”. The results here are like a single vowel, yet slightly longer, not as a sequence of two vowels. Nevertheless, it is definitely worthwhile to

⁵ Thanks to Akin Akinlabi for pointing this out.

compare instances of remnant vowels with clear cases of assimilation in VV sequences with both simple vowels and remnant vowels.

7. Summary and conclusion

While the standard account of vowel deletion in Yorùbá creates a phonological structure identical to a short vowel, new experimental results suggest that this should be revised. The duration of the vowel that remains after deletion in a VV sequence is slightly but significantly longer than a simple short vowel. A revised phonological account places the deletion process as a type of compensatory lengthening: the vowel root node deletes, and the remaining mora reassociates with the remnant vowel. However, the duration is only about 10% longer than a short vowel, which is much shorter than bimoraic vowels cross-linguistically. This could be related to the fact that Yorùbá does not have contrastive long vowels, so while there is phonological pressure for the vowel to lengthen, there is a low functional load. This work will hopefully lead to more research, both on the phonetics and the phonology, and on vowel deletion and related processes in Yorùbá.

Acknowledgements

Thanks to Adebola Isiah and Paul de Lacy for guidance and assistance with the experiment. Thanks to Katie Franich, Aaron Braver, Tadjou-N'Dine Mamadou Y. and three anonymous reviewers for their comments on this paper. Additionally, I would like to thank the participants of ACAL 48 at Indiana University where this work was presented as a poster. Lastly, thank you to Jeffrey Lamontagne for his skills with *ggplot*. Any remaining mistakes are my own.

References

- Ajíbóyè, O., R.-M. Déchaine, B. Gick and D. Pulleyblank. 2011. Disambiguating Yorùbá tones: At the interface between syntax, morphology, phonology and phonetics. *Lingua* 121(10): 1631–1648. <https://doi.org/10.1016/j.lingua.2011.05.008>
- Akinlabi, A. and F. Oyebade. 1987. Lexical and postlexical rule application: Vowel deletion in Yoruba. *Journal of West African Languages* 17(2): 23–42.
- Braver, A. 2013. Degrees of Incompleteness in Neutralization: Paradigm Uniformity in a Phonetics with Weighted Constraints. Unpublished PhD dissertation, Rutgers University.
- Broselow, E., S.-I. Chen and M. Huffman. 1997. Syllable weight: Convergence of phonology and phonetics. *Phonology* 14(1997): 47–82. <https://doi.org/10.1017/s095267579700331x>
- Clements, G.N. 1986. Compensatory lengthening and consonant gemination in Luganda. In L. Wetzels and E. Sezer (eds.) *Studies in compensatory lengthening*. Dordrecht: Foris. pp. 37–77. <https://doi.org/10.1515/9783110821666-004>
- De Chene, B. and S.R. Anderson. 1979. Compensatory lengthening. *Language* 55(3): 505–535. <https://doi.org/10.2307/413316>

- Flemming, E. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18(1): 7–44. <https://doi.org/10.1017/s0952675701004006>
- Francis, A.L., V. Ciocca and J.M.C. Yu. 2003. Accuracy and variability of acoustic measures of voicing onset. *The Journal of the Acoustical Society of America* 113(2): 1025–1032. <https://doi.org/10.1121/1.1536169>
- Gess, R. 2011. Compensatory lengthening. In M. van Oostendorp, C.J. Ewen, E. Hume and K. Rice (eds.) *The Blackwell companion to phonology* vol. 3. Malden, MA and Oxford: Wiley-Blackwell. pp. 1513–1536.
- Goldrick, M. 2000. *Turbid output representations and the unity of opacity*. Available online: <https://doi.org/10.7282/T3FJ2DSP> (Accessed 7 October 2019).
- Hayes, B. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2): 253–306.
- Klatt, D.H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59(5): 1208–1221. <https://doi.org/10.1121/1.380986>
- Ladefoged, P. 2003. *Phonetic data analysis*. Malden, MA: Blackwell.
- Mamadou Y., T.-N. 2017. Influences of Level Tones on Vowel Length in Yoruba. Unpublished MA project, University of Iowa.
- Ola Orie, O. and D. Pulleyblank. 2002. Yoruba vowel elision: Minimality effects. *Natural Language and Linguistic Theory* 20(1): 101–156. <https://doi.org/10.1023/a:1014266228375>
- Peterson, G.E. and I. Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32(6): 693–703. <https://doi.org/10.1121/1.1908183>
- Van Santen, J.P.H. 1992. Contextual effects on vowel duration. *Speech Communication* 11(6): 513–546. [https://doi.org/10.1016/0167-6393\(92\)90027-5](https://doi.org/10.1016/0167-6393(92)90027-5)