The acoustics of word-initial and word-internal voiced stops in Somali

Sabrina Bendjaballah

CNRS, Laboratoire de Linguistique LLING-UMR6310, France
sabrina.bendjaballah@univ-nantes.fr

David Le Gac

Université de Rouen-Normandie, Laboratoire Dylis-EA7474, France
david.legac@univ-rouen.fr

This article seeks to determine the acoustic correlates of gemination in Standard Somali (Afroasiatic, Cushitic), in particular whether closure duration is the primary acoustic correlate distinguishing singleton and geminate stops, with immediate consequences for the analysis of word-initial strengthening. We provide an acoustic analysis of word-initial and word-internal voiced singletons as well as of their geminate counterparts on the basis of a production experiment conducted with four native speakers. Three temporal and four non-temporal acoustic properties of /b d g/ and /bb dd gg/ are examined and systematically compared (closure duration, release burst duration, vowel duration; and closure amplitude, release amplitude, presence of a release burst, (de)voicing). We argue that the opposition between singleton and geminate voiced stops is primarily realized as the manner contrast approximant [β ð̞ ɣ] vs. stop [b d g]. Word-initially, Somali exhibits various peculiarities that are reminiscent of the cross-linguistically attested phenomenon of domain-initial strengthening. This article provides the first study of this phenomenon in Somali. We establish that word-initial /b d g/ and word-medial /bb dd gg/ share the same closure duration, release burst duration, and vowel duration within the Prosodic Word. They also have a similar closure amplitude, and voicing properties. Moreover, the acoustic properties of word-initial /b d g/ remain constant, and do not depend on their position in the prosodic hierarchy. On the basis of these results, the article also aims at providing new insights in the phonological representation of Somali geminates and word boundaries, and thus contributes to the understanding of word-initial strengthening in Somali.

1 Introduction

This study investigates the acoustic properties of word-initial and word-internal voiced stops in Somali, a Cushitic language of the Afroasiatic family. Like many other languages, Somali displays a phonological contrast between singleton and geminate consonants. Singleton consonants are attested in all contexts, while geminate consonants are attested in intervocalic position, only. The phonetic realizations of both singleton and geminate voiced stops have been the subject of very few studies (Armstrong 1934, Farnetani 1981, Barillot 2002). As a consequence, the acoustic correlates of gemination are still unclear at the present stage. In particular whether closure duration is the primary acoustic correlate distinguishing singleton and geminate stops remains an unsettled issue. A related unsolved issue concerns the phonetic realization of word-initial singleton voiced stops. These are reported by some authors (Armstrong 1934, Orwin 1994, Barillot 2002) to be geminated in some contexts. However, the empirical basis for this claim remains to be established. First, the context

in which this putative phenomenon obtains has not been precisely delineated. Second, the phonetic realization of word-initial singleton voiced stops seems to display a high level of variation. It is thus unclear whether Somali exhibits word-initial gemination, or rather a gradual phenomenon of domain-initial 'strengthening' correlated to the prosodic hierarchy, as widely attested cross-linguistically (see e.g. Fougeron & Keating 1997, Cho & Keating 2001, Keating et al. 2003, Keating 2006, Cho 2011), or neither of them.

This article provides an acoustic analysis of word-initial and word-internal /b d g/ as well as their geminate counterparts in Standard Somali. The analysis is based on a production experiment that was conducted with four Somali native speakers. A controlled corpus was designed to meet the following objectives: first, we aim at establishing the acoustic correlates of gemination in Somali; second, we aim at contributing to the understanding of word-initial gemination/strengthening in Somali. Three temporal and four non-temporal acoustic properties of word-internal /b d g/ vs. /bb dd gg/ are examined. They are compared with those of word-initial /b d g/ in three different contexts: nominal compounds, genitives and subject–object sequences. On the basis of the results of the experiment, the article also aims at providing new insights in the phonological representation of Somali geminates and word boundaries.

1.1 General background

Somali is a language spoken by 'about nine million people who occupy the north-eastern corner of Africa' (Saeed 1999: 1). In addition, there is an important number of Somali speaking communities in the diaspora. As a result, it is estimated that there are today approximately 20–25 million Somali speakers (Nilsson 2017). Somali belongs to the East-Cushitic branch of the Afroasiatic family. After Oromo, it is the Cushitic language with the largest number of speakers (Saeed 1999: 3). Somali has been written since the end of the 19th century. Different scripts have been used in the past, but today, Somali is uniformly written with a Latin-based orthography. The dialectal situation of Somali is not clearly understood at the present stage. Lamberti (1986), for instance, distinguishes 67 isoglosses, 17 varieties and five main dialects. However, there is a consensus that Somali dialects must be divided in three groups: Common/Northern Somali, Central Somali and Benadir (or Coastal) Somali (Saeed 1982, Abdullahi 2001).[1] The Somali Democratic Republic developed a strongly centralized language policy in the 1970s and 1980s, which led to the formation of a lingua franca, referred to as 'Standard Somali'. As pointed out in Nilsson (2018: 81),

> the variety spoken by the majority (often referred to as Northern Somali), was taken as the base, and the standard was formed as a certain compromise.

There is a certain degree of variation within this standard; however very little work has been done on this topic.

1.1.1 The consonant inventory of Standard Somali and the distribution of singleton and geminate consonants

The consonant inventory of Standard Somali as given, among others, by Armstrong (1934), Cardona (1981), Orwin (1995), and Saeed (1999) is reproduced in Table 1. Where Somali orthography diverges from IPA, we give the transcription in Somali orthography in parentheses. Grey cells indicate the consonants that may geminate.

Singleton consonants are attested in all positions: word-initially, word-finally, as well as word-internally, in onset and coda position. Geminates are attested in intervocalic context,

---

[1] Relevant work on Somali dialectology includes Ehret & Ali (1984), Lamberti (1986), Abdullahi (2001) and Abdirachid (2011); see Abdullahi (2001: Chapter 3) for an overview of the literature. Ehret & Ali (1984) is exclusively based on lexicostatistics. Lamberti (1986), Abdullahi (2001) and Abdirachid (2011) describe the variation displayed at the lexical level as well as that observed at the phonetic, phonological, morphological and syntactic levels.

Table 1 The consonant inventory of Standard Somali.

		Labial	Dental	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Laryngeal
Stops	[–voi]		t		t͡ʃ (j)	k	q		ʔ (')
	[+voi]	b	d	ɖ (dh)		ɡ			
Fricatives	[–voi]	f	s		ʃ (sh)	x (kh)		ħ (x)	h
	[+voi]							ʕ (c)	
Nasals	([+voi])	m	n						
Approximants	([+voi])	w	l r		j (y)				

only. They are either lexically given (e.g. /ʕiddale/ 'owner'), or the result of assimilation rules (e.g. /mindi/ → /middi/ 'knife'). At the lexical level, there are no word-initial or word-final geminates. It is generally considered that only a subset of Somali consonants has geminate counterparts at the phonetic level: b, d, ɖ, ɡ, m, n, l, r. All authors report that geminate sonorants are consistently longer than their singleton counterparts. Duration thus appears to be the main phonetic correlate that distinguishes geminates and singletons in this class of Somali consonants. In this respect, Somali behaves like the vast majority of the world's languages (see e.g. Ridouane 2010, Hamzah, Fletcher & Hajek 2016 for an overview). For the stops, the situation is less clear-cut. /ɖ/ has many idiosyncratic peculiarities. In particular, it is reported to display a wide range of different realizations. As reported by Armstrong (1934: 121–122), it may be articulated with an implosive quality, involving the contraction of the pharynx and the raising of the larynx, followed by the relaxation of the pharyngeal contraction, and the lowering of the larynx. In intervocalic position, it may also be realized as a flapped [ɾ] (Armstrong 1934: 122). Given the complexity involved in the realizations of /ɖ/, which are independent of our issue, we consider that this segment merits being treated separately, and we will concentrate on /b d ɡ/. For this class of voiced stops, the facts are not clear: it has been suggested that duration is a consistent correlate of gemination, but other authors, e.g. Barillot (2002), challenge this assumption, and suggest that the manner of articulation should rather be considered the primary correlate of gemination.

1.1.2 Tonal accent and prosodic constituency

Since the seminal study of Hyman (1981), Somali is generally considered a tonal- or pitch-accent language. Tonal accent (TA) consists of a phonological high tone, realized as a high or mid pitch target. This pitch target is associated with an intensity peak, which is referred to either as 'stress' (e.g. Armstrong 1934, Andrzejewski 1964, Orwin 1995) or 'accent' (Hyman 1981; Banti 1988; Saeed 1999; Le Gac 2001, 2003a, b among others). Syllable duration is not a phonetic correlate of TA. The TA-bearing unit is the vocalic mora. There is at most one TA per word, which occurs either on the penultimate or on the last mora of the word. TA is not lexically distinctive but determined by various grammatical features such as gender, number, case, verb inflection etc. (Andrzejewski 1964, 1979; Hyman 1981; Banti 1988; Orwin 1995; Saeed 1999; Le Gac 2001, 2003a, b).

In Somali, as in all languages, a correct identification of the prosodic constituency relies both on segmental and suprasegmental features. Segmental diagnosis tools include sandhi rules. The lenition of intervocalic stops, for instance, has been shown to apply within the Prosodic Word (ω), e.g. /t/ → [ð]: (magaaló+ta)ω 'town+THE' → [maga:láða] 'the town' vs. (lá)ω (tág)ω 'with go' → [láˠtʰæɡˠ] 'go with (it/him/her)!'[2] As for the suprasegmental

---

[2] Except if otherwise specified, we use the Standard Somali orthography throughout this article, and transcribe TA by an acute accent on the TA-bearing vocalic mora. Prosodic constituent boundaries are marked by parentheses.

features, since there is one, and only one, TA per word, TA is assumed to be a diagnosis of Prosodic Wordhood (Hyman 1981; Le Gac 2001, 2003a, b; Green & Morrison 2016; Downing & Nilsson 2019).[3] An independent noun, e.g. *baabúur* 'car', with a single TA therefore constitutes a Prosodic Word: $(\acute{N})_\omega$. A nominal compound, e.g. *cilmi-baarís* 'science-research' → $(cilmi\text{-}baarís)_\omega$ 'scientific research', which has a single TA located either on the penultimate or on the last mora of the second noun, also constitutes a single Prosodic Word: $(N1\ \acute{N}2)_\omega$. By contrast, in certain nominal phrases like indefinite genitive constructions, each noun has its own TA, e.g. *batéri baabuúr* 'battery car' → $(batéri)_\omega$ $(baabuúr)_\omega$ 'a battery of a car'. In these structures, each noun is associated with its own ω: $(\acute{N}1)_\omega\ (\acute{N}2)_\omega$.[4]

Despite recent important work, the prosodic structure of Somali remains largely understudied. This is particularly the case of the levels located above ω. As a consequence, the number of prosodic constituents above ω and their defining criteria are still not clear. However, certain processes have been shown to be a diagnosis of the Phonological Phrase (ϕ).[5] More specifically, Le Gac (2001, 2018) proposes that the domain of application of downdrift (or downstep) is ϕ.[6] In addition, he notes that ϕ generally ends up with a low or high edge tone, and it may be followed by a pause. These diagnosis tools make it possible to establish for instance that a Noun1 Noun2 indefinite genitive construction constitutes a ϕ: in this construction, the tonal accent of Noun2 is usually pronounced a bit lower than that of Noun1 (Hyman 1981, Le Gac 2001); no edge tone and no pause intervene between Noun1 and Noun2. By contrast, no downdrift applies in sequences of two syntactically independent nouns, e.g. a subject Noun1 followed by an object Noun2, or an adverb Noun1 followed by a subject Noun2. In these configurations, Noun1 and Noun2 normally end up with an edge low or high tone, and they may be followed by a pause. Each noun thus constitutes a prosodic constituent equivalent to ϕ. At a higher level, these sequences may be grouped together into higher prosodic constituents like the Intermediate Phrase or the Intonational Phrase, which end up with a low boundary tone L% (Le Gac 2001).

To sum up, we will assume the prosodic structures in (1) for the four constructions that are relevant in our experiment. Simple nouns constitute a ω (1a). Nominal compounds are grouped together within a single ω (1b). In that sense, they have the same structure as simple nouns. In indefinite genitive constructions (1c), N1 and N2 form two ωs inserted into a unique ϕ. Finally, in subject–object sequences (1d), N1 and N2 constitute two ωs, which in turn constitute each an independent ϕ.

(1) a. Simple nouns: $(\acute{N})_\omega$

    b. Nominal compounds: $(N1\ \acute{N}2)_\omega$

    c. Indefinite genitive constructions: $((\acute{N}1)_\omega\ (\acute{N}2)_\omega)_\phi$

    d. Subject–object sequences: $((\acute{N}1)_\omega)_\phi\ ((\acute{N}2)_\omega)_\phi$

---

[3] Within a framework that admits the recursion of prosodic constituents, Green & Morrison (2016) assume that the constraint 'a single TA per prosodic word' applies at the 'Prosodic Word Max' level, whereas Downing & Nilsson (2019) assume it to apply at the 'Prosodic Word Min' level.

[4] In addition, nominal compounds differ from genitive constructions at the syntactic level: nominal compounds are marked by a single determiner at the right edge of the compound (*cilmi-baaris+ta* 'the scientific research' vs. *bateri+ga baabuur+ka* 'the battery of the car') and no element can be inserted between the two nouns of a compound.

[5] See Le Gac (2001, 2003a, b, 2018), Green & Morrison (2016), Downing & Nilsson (2019) for some phonological properties of the Noun Phrases located at the level of the Phonological Phrase or above.

[6] Le Gac (2001) uses the term 'Groupe Prosodique' to refer to the domain of application of downdrift. The 'Groupe Prosodique' is a prosodic constituent that is located between the Prosodic Word and the Intermediate Phrase. In this sense, it is equivalent to the Phonological Phrase.

1.2 The phonetic correlates of gemination in Somali: Closure duration only?

To our knowledge, there are only three studies that investigate the phonetic realization of singleton and geminate consonants in Somali in detail: Armstrong (1934), Farnetani (1981) and Barillot (2002).[7] The results obtained by Armstrong (1934) as well as the acoustic analysis conducted by Farnetani (1981) are coherent with what has been observed in many languages: the contrast between geminates and singletons primarily relies on duration. Barillot (2002) questions this assumption: he claims that the manner of articulation (spirantized consonant vs. stop) is the primary contrast that opposes singleton and geminate stops. Before we proceed, note that this state of affairs can hardly be ascribed to dialectal or sociolinguistic variation: the speakers recorded by Armstrong and Farnetani are of a different geographical origin, and separated by more than one generation. However, they display similar realizations. By contrast, the speakers consulted by Farnetani and Barillot are originally from the same area; yet, their realizations significantly diverge.

Armstrong (1934) represents a landmark in Somali studies in that it is the first detailed analysis of Somali phonetics. Her work is based both on an auditory approach and on kymograph tracings. Two speakers have been consulted, both from the North of Somalia. The syntactic and phonological environments have not been controlled. Armstrong (1934: 117) notes that the 'length of both consonant and vowel sounds is important and often significant'. The realization of the contrast singleton vs. geminate voiced stop involves a length contrast: /bb dd gg/ are realized as 'double'[8] /b d g/. In addition, Armstrong notes that, in some cases, /b d g/ are spirantized to [β ð ɣ]. However, there is inter-speaker variation: spirantization obtains for one speaker, but not for the other. She also suggests that spirantization takes place especially, but not always, after a stressed vowel. Regarding /bb dd gg/, Armstrong observes various degrees of voicing, depending again on the speaker. Armstrong's findings for singleton and geminate /b d g/ in intervocalic position are summed up in Table 2.

Table 2 Armstrong (1934): Intervocalic realizations of singleton and geminate voiced stops.

Intervocalic /b d g/	Intervocalic /bb dd gg/
[β ð ɣ] (especially after a stressed syllable) or [b d g] inter-speaker variation	'double' b d g fully or partially voiced inter-speaker variation

Farnetani (1981) represents the only comprehensive acoustic study of the Somali segmental system. Broadly speaking, she confirms Armstrong's findings. Her corpus consists of a list of 90 utterances ranging from words in isolation to complete sentences produced by four male speakers aged between 34 years and 41 years. Farnetani (1981) examines the contrast between intervocalic /b d g/ and /bb dd gg/ by investigating the three following phonetic characteristics: closure duration, manner of articulation and voicing. Her results are summarized in Table 3. The average closure duration of /bb dd gg/ is twice or three times longer than that of /b d g/ (157–174 ms vs. 49–71 ms respectively). All realizations of /bb dd gg/ are characterized by a complete closure of the articulators, i.e. they are 'true' stops,

---

[7] The topic is briefly addressed in Hassan (1994), Orwin (1994), Saeed (1999), and Edmonson, Esling & Harris (2003). These studies (except for Hassan 1994) all rely on Armstrong's data, or take over her findings.

[8] This is the term used by Armstrong to refer to length (see Armstrong 1934: 119, 120ff.). Armstrong does not seem to have used her kymograph tracings to measure consonant duration. Rather, she seems to base her evaluation of 'double' consonants on her subjective impressions. Kymograph tracings seem to have been used in order to establish whether a given sound is voiced or not, only (Armstrong 1934: 122, 127). This observation prompted us to conduct a new, controlled, acoustic study of Somali stops.

Table 3 Farnetani (1981): Acoustic properties of intervocalic /b d g/ and /bb dd gg/. Voicing duration is measured as % over closure duration; n = number of items.

	Intervocalic Singletons (vCv)				Intervocalic Geminates (vCCv)			
	n	Closure duration (ms)	Voicing duration	Manner	n	Closure duration (ms)	Voicing duration	Manner
b	22	50	100%	all approximant	15	160	9: 100% 6: 80%	all stop
d	14	49	100%	10: approximant 4: stop	16	174	11: 100% 5: 78%	all stop
g	13	71	100%	all approximant	8	157	5: 100% 3: 70%	all stop

Table 4 Barillot (2002): The realization of intervocalic singleton and geminate voiced stops.

Intervocalic /b d g/	Intervocalic /bb dd gg/
[β ð ɣ]	[bb dd gg] or [b d g]

whereas /b d g/ are realized as approximants in 88% of the cases under study.[9] Finally, /b d g/ are always fully voiced, whereas /bb dd gg/ are partially devoiced in 36% of the cases under study.

Against this background, Barillot (2002) offers a somewhat different picture of the distribution of continuants and stops in Somali. His corpus consists of words in isolation elicited with one native speaker, and his results are based on a qualitative and auditory approach. Barillot (2002: 223–226) notes that intervocalic voiced stops surface in three shapes: singleton voiced fricatives, singleton voiced stops and geminate voiced stops. As can be seen in Table 4, Barillot analyses the situation as follows: intervocalic /b d g/ always surface as voiced fricatives, and intervocalic /bb dd gg/ surface either as geminate or singleton voiced stops. For /bb dd gg/, the choice between one of the two realizations does not seem to be predictable. This constitutes a major difference with Armstrong (1934) and Farnetani (1981), for which geminates are ALWAYS realized as long segments.

We are now in a position to evaluate the main issues that arise from the literature on Somali word-internal voiced stops. The first one concerns the singletons. The literature suggests that they tend to be 'lenited', in particular via spirantization. Our aim is first to verify whether, and to which extent, singleton voiced stops are lenited in Standard Somali, and second, to define the acoustic properties and the phonetic category of the 'lenited' segments: are word-internal voiced singletons realized as 'weakened' stops (with very short closure and/or release duration), or rather as approximants? The second issue pertains to the geminates. It is unclear whether closure duration constitutes the primary correlate of gemination in Somali. This inconsistency is likely to be due to the fact that other acoustic parameters override closure duration. Indeed, various articulatory and acoustic parameters have been reported to contribute to the perceptual effect of gemination cross-linguistically.

There is broad agreement about the fact that closure duration plays a major role in distinguishing singletons and geminates cross-linguistically. In the surveys of 24 and 39 languages provided in Ridouane (2010) and Hamzah et al. (2016) respectively, the contrast between word-medial singletons and geminates consistently involves a length contrast (with a longer closure duration for the geminates). The other acoustic attributes less consistently oppose singletons and geminates. Two temporal parameters however stand out, and need to be carefully

---

[9] Farnetani defines 'approximants' as 'continuants without frication noise'. They may have formants that are strongly attenuated, but remain visible.

considered: vowel duration and release duration. In eight of the 24 languages considered in Ridouane (2010), vowel duration has been found to be shorter before geminates than before singletons. These include unrelated languages like Austronesian languages (Cohn, Ham & Podesva 1999), Bengali (Lahiri & Hankamer 1988), and Tashlhiyt Berber (Ridouane 2007). This pattern however is far from being systematic: in Japanese for instance, vowel duration directly covaries with consonant duration (Kingston et al. 2009). Positive VOT or release duration contributes to the acoustic difference between geminates and singletons in some languages like Tashlhiyt Berber, where the release duration of geminate voiced stops is significantly longer than that of singleton voiced stops (Ridouane 2007). Non-temporal characteristics have been less widely investigated, but they have also been reported to be involved in the contrast between singletons and geminates. Release amplitude has been shown to be higher in geminates than in singletons. In particular, the release burst is reported to be produced with significantly greater energy in geminate stops than in singleton stops (Hamzah, Fletcher & Hajek 2012 for Kelantan Malay, and Ridouane 2007 for Tashlhiyt Berber). In Tashlhiyt Berber, singleton voiced stops are sometimes characterized by the absence of a release burst altogether, a state of affairs that enhances the contrast between singletons and geminates: geminates show a burst vs. singletons do not. (De)voicing is an additional relevant parameter. Recall that Farnetani (1981) reports that geminate stops are partially devoiced in 36% of her data. Indeed partial, or total, devoicing of geminate voiced stops has been reported to enhance the contrast between singleton and geminate voiced stops in various languages (see for instance Ohala 1983 for More, Ridouane 2007 for Tashlhiyt Berber, and Jaeger 1978: 322 for additional cases). Finally, the spectral characteristics of the consonants and the surrounding vowels are argued to be involved in the contrast between singleton and geminate consonants, e.g. in Malayalam (Local & Simpson 1999).

In this article, we will assess to which extent the relevant temporal and non-temporal parameters mentioned above (viz. closure duration, release duration, vowel duration; and closure amplitude, release amplitude, presence/absence of release burst, devoicing respectively) are involved in the contrast opposing singleton and geminate voiced stops in Standard Somali.[10] On this basis, we will establish whether closure duration is the primary correlate of gemination, or whether it is overridden by other acoustic correlates.

1.3 The phonetic realization of word-initial voiced stops: Word-initial gemination or domain-initial strengthening?

Armstrong (1934), Farnetani (1981), Orwin (1994), and Barillot (2002) observe a positional asymmetry: word-initially, singleton stops seem to be longer than word-internally. However, neither the contexts in which this phenomenon obtains nor its precise phonetic characteristics have been clearly defined yet.

According to Armstrong (1934: 119–123), word-initial /b d g/ are realized with not much voice and without aspiration. She draws attention to the fact that in connected speech 'double consonants', and in particular voiced plosives, frequently occur word-initially (Armstrong 1934: 138–139). Armstrong's findings for the phonetic realizations of word-initial /b d g/ are summed up in Table 5.

Armstrong suggests that this 'doubling' may be due to the presence of 'a stressed syllable (ending in a short vowel and pronounced usually) with the high-level tone' immediately

---

[10] Somali does not straightforwardly lend itself to an examination of the spectral properties of the surrounding vowels. Indeed, Somali has two series of vowels that are usually defined as front (= [+ATR]) and back (= [−ATR]). These series are involved in complex vowel harmony processes, which apply in sequences that are longer than the word (Armstrong 1934, Andrzejewski 1955, Saeed 1999). Since vowel harmony has not been experimentally investigated at the present stage, we feel that we are not in a position to safely control this parameter, and infer generalizations from the properties of the vowels preceding and following singletons and geminates in Somali.

Table 5 Armstrong ([1934](#)): Phonetic realizations of word-initial /b d g/.

Word-initial /b d g/
[b]/[b̥] [d]/[d̥] [g]/[g̥]
with not much voice or not at all
'double' b d g

Table 6 Farnetani ([1981](#): 69–70): Acoustic properties of word-initial /b d g/ after a word in the sentence (x#Cv) and in isolation (##Cv). Closure duration is not given in ##Cv because the beginning of the closure cannot be acoustically determined in this context; n = number of items.

	x#Cv			##Cv			
	n	Closure duration (ms)	Manner	n	Closure duration	Voicing duration (ms)	Manner
b	8	92	stop	22	–	83	stop
d	3	117	stop	16	–	90	stop
g	7	134	stop	8	–	90	stop

before the consonant, e.g. *kú ḍḍɛh* 'say to him/her' (Armstrong [1934](#): 139). Orwin ([1994](#): 59ff.) adopts this conclusion, and considers that word-initial geminates occur after stressed syllables, in specific syntactic positions; in this sense, Somali seems to display a phenomenon that is comparable to the Italian *raddoppiamento sintattico*. More generally, this would suggest that stress in Somali has a strengthening effect similar to that reported in various languages (see e.g. Hirst & Di Cristo [1998](#), Turk & White [1999](#), Cho & Keating [2009](#)). The correlation between stress and word-initial geminates or 'double' consonants in Somali should however be taken cautiously. Indeed, stress seems to have contradictory effects: recall from Section [1.2](#) that, according to Armstrong, stress is involved in the LENITION of intervocalic /b d g/ to [β ð ɣ]. This observation questions the role of stress as a trigger for word-initial gemination/'doubling'.

Farnetani ([1981](#): 69–70) confirms a clear contrast in the realization of word-initial vs. word-internal /b d g/. Word-initial /b d g/ are realized as stops and show a certain level of devoicing in absolute initial position, while they are always fully voiced in connected speech. In addition, word-initial /b d g/ are realized up to twice as long as their counterparts in word-internal intervocalic context. However, the average closure duration of word-initial /b d g/ clearly does not reach that of intervocalic /bb dd gg/ (92–134 ms vs. 157–174 ms), see Table [6](#). Stress does not seem to play any role in Farnetani's data.

The word-initial position is known to be a strong position in phonology: in many languages, word-initial consonants are strengthened, both in synchrony and in diachrony.[11] Recent work in phonetics reveals a temporal and spatial expansion of several articulatory and acoustic parameters in this context (e.g. the quantity of linguo-palate contact, closure seal duration, and VOT). This temporal and spatial expansion is often referred to as (domain-) initial strengthening. Domain-initial strengthening has been reported to be cross-linguistically proportional to the position of the segment in the prosodic hierarchy: the higher up the segment is located, the more it is strengthened. For instance, a segment located at the beginning of an Intonational Phrase will be more strengthened than a segment located at the

---

[11] See among many others, Ségéral & Scheer ([2001](#)) for facts and analysis, and Hock ([1991](#)) for historical sound changes involving the strengthening of a word-initial approximant to an obstruent, or a continuant to a non-continuant consonant.

beginning of a Phonological Phrase or a Prosodic Word (see e.g. Fougeron & Keating 1997, Cho & Keating 2001, Keating et al. 2003, Keating 2006, and Cho 2011 for a review).

The fact that Somali word-internal and word-initial singleton stops have different realizations, as evidenced by Armstrong (1934) and Farnetani (1981), might be ascribed to a phenomenon of domain-initial strengthening. Under this hypothesis, initial strengthening would be marked by the fact that the voiced stops surface as 'true' stops, with a clear release burst and a longer closure duration. This hypothesis might also provide an explanation for Armstrong's mention of initial 'double' consonants: 'doubling' would result from the presence of an important prosodic boundary to the left of the consonant (and not from stress). This hypothesis makes a clear prediction on the contexts in which consonant 'doubling'/gemination is expected to take place: the higher up in the prosodic structure the segment is located, the more it is strengthened. However, this prediction is not immediately borne out. Indeed, Barillot (2002: 134) points out a puzzling fact concerning the realization of the initial consonant of Noun2 in Noun1 Noun2 compounds: according to standard dictionaries (e.g. Zorc & Osman 1993), this consonant sometimes surfaces as a geminate.[12]

(2) a.   hánti + goosád   >   hanti**gg**oosád
         property + cutting for oneself   capitalism
     b.   gaári + daméer   >   gaari**dd**améer
         car + donkey   car pulled by a donkey

As shown in Section 1.1.2, Somali nominal compounds constitute a single Prosodic Word: they bear a single tonal accent (and not two), they are marked by a single determiner at the right edge of the compound, no element can be inserted between the two nouns, etc. Under the hypothesis of a classical phenomenon of domain-initial strengthening, the initial consonant of Noun2 in nominal compounds is not expected to geminate. Indeed, word-internal boundaries are more deeply embedded in the prosodic hierarchy than boundaries between constituents. Gemination is thus expected to occur between two independent syntactic constituents, e.g. between a subject Noun Phrase and an object Noun Phrase (which each form a Phonological Phrase or a higher constituent, see Section 1.1.2), rather than between the two nouns of a nominal compound. However, no gemination has been reported between two syntactic constituents.

In this study, we thus aim at answering the following question: Are Somali word-initial voiced stops realized as true geminates, or do they undergo domain-initial strengthening? Or neither: do /b d g/ simply have specific realizations in word-initial position? If word-initial voiced stops are realized as true geminates, we expect them to share at least one acoustic characteristic with (word-internal) lexical geminates. In particular, we expect word-initial singletons to share the primary acoustic correlate of lexical geminates, e.g. they should display similar values for their closure duration. If Somali exhibits domain-initial strengthening, we expect at least one acoustic characteristic of word-initial voiced stops to gradually increase with the level in the prosodic hierarchy. This could be the case of the primary correlate of gemination (e.g. closure duration) and/or of other acoustic parameters (such as release duration and/or amplitude). Beyond a given threshold, this increase would give the impression that gemination takes place. Finally, it could be the case that word-initial stops are neither geminates, nor subject to domain-initial strengthening, but simply characterized by one or

---

[12] This gemination cannot be ascribed to word prominence since it occurs in cases where word prominence is located at a distance. It does not result from the rules of Somali phonotactics either, since Somali sandhi rules trigger lenition of intervocalic stops. Finally, note that there is no reason to assume the presence of a morpheme between Noun1 and Noun2 that would be assimilated by the initial consonant of Noun2 and thus trigger gemination. Such a morpheme never surfaces if Noun1 is a consonant-final noun, e.g. *dáb-ḏul* (fire-earth) 'impetigo, chickenpox', see Barillot (2002: 133ff.) for discussion.

more acoustic correlates that are absent in word-internal position. For instance, word-initial consonants would be systematically realized with a release burst while word-medial consonants never would. The presence vs. absence of a release would give rise to the auditive impression of a contrast between geminates and singletons.

In this article, we will evaluate these options. We will gradually vary the level of the prosodic boundary located before the word-initial test stop, and systematically compare its acoustic correlates with those of lexical geminates and word-internal singletons. This will make it possible to determine whether word-initial /b d g/ are realized as geminates, or undergo domain-initial strengthening (or neither).

More generally, this study is meant as a contribution to the under-investigated field of Somali phonetics, with a view to offer new experimental insights into the acoustic correlates of gemination and domain-initial strengthening in Somali, and beyond.

2 Method

2.1 Corpus

A production experiment relying on a controlled corpus was designed to establish which acoustic correlates distinguish singleton /b d g/ from their geminate counterparts, and whether word-initial /b d g/ are realized as geminates or undergo domain-initial strengthening. Singleton and geminate /b d g/ are examined in intervocalic position in morphologically simplex nouns (3a). In addition, we consider Noun1 Noun2 sequences, in which the test stop is the initial consonant of Noun2 (3b). We investigate the three following contexts: (i) adjacent nouns that form a nominal compound together, (ii) adjacent nouns that form a constituent together, more specifically indefinite genitive constructions, and (iii) adjacent nouns that do not form a constituent, more specifically, sequences of a subject followed by a direct object. These three contexts were selected because the initial stop of Noun2 is expected to be embedded at a different level of the prosodic structure in each of them, namely (i) within the Prosodic Word (ω) in nominal compounds, (ii) at the boundary between two Prosodic Words within a Phonological Phrase (ϕ) in genitive constructions, and (iii) at the boundary between two Phonological Phrases in subject–object sequences (see Section 1.1.2). To sum up, /b d g/ were inserted in the following five conditions:

(3)   CONTEXT   ABBREVIATION

    a. Word-internal

      morphologically simplex noun (Ń)$_\omega$ singleton   LexC
                                                                 geminate   LexCC

    b. Word-initial

      N1 N2 compound   (N1 Ń2)$_\omega$ singleton   Cmp
      N1 N2 indefinite genitive   ((Ń1)$_\omega$ (Ń2)$_\omega$)$_\phi$ singleton   Gen
      N1 N2 subject–object   ((Ń1)$_\omega$)$_\phi$ ((Ń2)$_\omega$)$_\phi$ singleton   Ind

The test stop was inserted in carrier sentences with the following structure: [X (N1) N2 *waxaa* Verbal Complex Y]. 'Verbal Complex' refers to the verb, potentially preceded by different particles, and 'X' and 'Y' correspond to any Noun Phrase or temporal/locative Adverbial Phrase. Focus is known to have an important influence on prosodic phrasing in many languages, no matter whether they are based on a stress, pitch-accent or tone system (for a review, see e.g. Gussenhoven 2004, Ladd 2008). The only systematic experimental study on this effect in Somali is the one conducted by Le Gac (2001, 2003a, b). It suggests

Table 7 Number of sentences.

Context		b	d	g	Total
LexC	(word-internal singleton stop)	6	6	6	18
LexCC	(word-internal geminate stop)	5	5	2	12
Cmp	(initial singleton stop of N2 in N1 N2 compounds)	5	5	3	13
Gen	(initial singleton stop of N2 in N1 N2 genitives)	7	7	10	24
Ind	(initial singleton stop of N2 in N1 N2 subject–object sequences)	5	5	6	16
Total		28	28	27	83

that focus, and especially contrastive focus, does have an influence on Somali prosody: it involves the insertion of specific boundary tones, pitch resetting on the focused noun, and pitch range-compression on post-focused elements. In order to exclude any potential focal prominence or prosodic rephrasing in the environment of the test stop, we used *waxaa*-constructions in our experiment. *Waxaa* [waħaː] is one of the Somali focus particles that focuses the LAST constituent of the sentence, i.e. Y in our carrier sentences (Puglielli 1984, Lecarme 1999, Saeed 1999). This ensures that neither X, nor N1, nor N2 is under focus.

In all sentences used in the experiment, the serial position of the test segment within the sentence was kept constant. Indeed, some authors have argued that articulatory declination may modify the articulation of a given segment because of its early-to-late position in the sentence. Since the notion of articulatory declination is controversial (see Krakow, Bell-Berti & Wang 1994 vs. Fougeron & Keating 1997), we excluded any potential effect of this sort: all test consonants were located at the onset of the sixth syllable.

Finally, the experiment also needed to exclude a potential influence of tonal accent. Indeed, as mentioned in Section 1, word prominence is likely to have an important, but unclear, influence on the realization of the consonants. We controlled this factor by ensuring that the syllables immediately preceding and following the target consonants did not bear tonal accent: all test consonants were preceded by unaccented /i/ and followed by unaccented /a/.[13]

Based on these principles, we drafted a corpus that was subsequently amended by three consultants, who were not aware of the purpose of the experiment: one linguist expert in Somali and two native speakers of Somali. They were asked to check the grammaticality of the sentences, as well as the appropriateness of the selected lexical items. In particular, they were asked to exclude expressions that could be specific to a particular dialect. As a result, the corpus used in the experiment consists of 83 sentences instantiating /b d g/ in the five relevant contexts, and distributed as shown in Table 7.

The number of sentences including indefinite genitive constructions is higher than the number of sentences in the other categories (Gen = 24). This was done on purpose. Indeed, the semantic divide between indefinite genitives and nominal compounds is not clear: indefinite genitives may be interpreted and realized as compounds.[14] We thus increased the number of indefinite genitive constructions in order to have enough data for this condition in case an intended genitive construction was interpreted as a compound by the speakers. At another

---

[13] Final /i/ in Noun1 was chosen for two reasons: first, final /o e/ sometimes alternate with /a/ before consonant and involve a final morpheme or a latent consonant (Cardona & Agostini 1981, Barillot 2002); second, final /a u/ are rare in lexical categories (they mostly appear as inflectional markers). Post-consonantal /a/ was chosen because it is more common than the other vowels and thus facilitated the design of the corpus. The length of the /i/ preceding the test consonant was controlled (always short); however, the length of the /a/ following the test consonant could not always be kept constant.

[14] Recall from Section 1.1.2 that the two structures may unambiguously be distinguished on the basis of their word prominence: indefinite genitives have two instances of tonal accent whereas compounds have only one.

level, the data are scarce for the following two configurations: /g/ as a word-internal geminate (LexCC) and /g/ as the initial consonant of Noun2 in compounds (Cmp). This is due to the fact that our consultants uniformly accepted only two nouns with intervocalic /gg/ (*higgaad* 'orthography', and *miigganaan* 'goodness') and three compounds with /g/ as the initial segment of Noun2 (*cabsi-gal* 'panic', *caqli-gaabyo* 'unintelligent persons', and *hanti-goosato* 'capitalists').[15]

The sentences used in the five conditions are exemplified for the test segment /b/ in (4).

(4) Examples of carrier sentences used for /b/ (orthographic transcription; FOC = focus particle)

    a. *Word-internal singleton stop (LexC)*

        Sannad ka hor li**b**aaxyo waxay dileen Diiriye.

        a year ago   lions   FOC   killed Diiriye

        'Lions killed Diiriye a year ago.'

    b. *Word-internal geminate stop (LexCC)*

        Abuubakar dhi**bb**aannadiisii wuxuu ku   daaweeyey jeermitire.

        Abubakar wounds.HIS   FOC   with   treated   disinfectant

        'Abubakar treated his wounds with disinfectant.'

    c. *Initial singleton stop of N2 in N1 N2 compounds (Cmp)*

        Wasiirku   cilmi-**b**aaris   wuxuu ku bixiyay siddeed malyuun oo doollar.

        minister.THE science-research FOC   on spent   eight million dollars

        'The minister spent eight million dollars on scientific research.'

    d. *Initial singleton stop of N2 in N1 N2 genitives (Gen)*

        Saaka   bateri **b**aabuur waxaan ka   soo iibasaday Kulmiye.

        this morning battery car   FOC   from   bought   Kulmiye

        'This morning I bought a car battery from Kulmiye.'

    e. *Initial singleton stop of N2 in N1 N2 subject–object sequences (Ind)*

        Berrito   Cali **B**atuulo wuxuu bari   doonaa Gadiid.

        tomorrow Ali Batuulo FOC   introduce will   Gadiid

        'Tomorrow Ali will introduce Gadiid to Batuulo.'

2.2 Subjects and procedure

Five Somali native speakers were recorded in March 2019 in London (UK): four male speakers (CNA, CQA, CRX and MAX) and one female speaker (DEE), all aged between 43 and 50. They were all born and raised in Somalia, and lived in London at the time of recording. All of them claimed to use both Somali and English on a daily basis. The data produced by MAX

---

[15] Given the scarcity of the data, we included one token with /o/ instead of /a/ after the test stop.

Table 8 Speakers recorded for the study.

Speaker	Age (gender)	Birthplace (Region)	Primary & secondary school	University	Profession
CQA	50 (M)	Buuloburde (Hiiraan)	Buuloburde (Hiiraan)	Mogadishu; Malaysia	Somali teacher
CNA	43 (M)	Mogadishu	Banaadir (Mogadishu)	London	Librarian
DEE	49 (F)	Mogadishu	Mogadishu	–	Housewife
CRX	44 (M)	Beledweyne (Hiiraan)	Beledweyne (Hiiraan); Mogadishu	UK	Advisor in capacity building for the government in Somalia

had to be excluded from the analysis because of disfluencies in reading and the insertion of many pauses. Consequently, we report the results obtained for four Somali native speakers: CNA, CQA, CRX and DEE. Two of them (CQA and CRX) come from the two major towns of the Hiiraan district in the central part of Somalia, where they attended both primary and secondary school. The two other speakers (CNA and DEE) come from the area around Mogadishu, where they attended primary and secondary school. Geographically speaking, Hiiraan is the closest region north from Mogadishu, and Hiiraan and Mogadishu are part of the same dialectal group (Lamberti 1986; Abdirachid 2011: 496). The four speakers included in this study thus constitute a homogeneous group, and the data recorded in the experiment are representative of the variant of Standard Somali spoken in the area. The relevant information on the respective background of the speakers appears in Table 8.

The speakers were recorded under the same conditions in one recording session each. The sessions took place in the recording studio of the School of Oriental and African Studies, using a high-quality electret condenser microphone (Audio-technica AT4033) and a digital recorder Marantz PMD671. The recordings were digitized in the WAV format at 44100 Hz and 24-bit.

None of the speakers was aware of the aim of the experiment. Each sentence of the corpus was transcribed in the standard Somali orthography, and printed on a specific sheet, yielding 83 different sheets. These sheets were randomized and presented by the experimenter to the speakers one by one. The speakers could interact with the experimenter in order to check the meaning of the intended sentence, and exclude a wrong interpretation. The speakers were asked to first review the entire sentence, and then to produce it in the most natural way, avoiding the insertion of unnatural breaks. The speakers produced the sentences one by one. The lapse of time between two sentences was controlled by the experimenter. When the speaker was done with all sheets, s/he was exposed to them again, but in the reverse order, thus starting with the last item of the first series, and producing the sentences again until the first item of the first series was reached. Finally, s/he was asked to produce again all sentences in the original order. As a consequence, each sentence was produced and recorded at least three times. The number of repetitions was chosen so as to ensure at least 15 recordings for each consonant in each condition. If there were not enough distinct sentences (i.e. /g/ as a word-internal geminate, and in nominal compounds), the speaker produced the sentences as many times as necessary to obtain at least 15 tokens. Productions with hesitations or restart were discarded, and the speaker was asked to produce the sentence again.

2.3 Labelling procedure and measures

The recorded sentences were labelled and analysed with Praat (Boersma & Weenink 2019). The acoustic analysis and the labelling procedure were conducted using the broad-band spectrograms and the corresponding waveforms of the utterances. Three temporal and four

non-temporal parameters were considered (see (5) and (6) below). These parameters correspond to the acoustic correlates that have been reported to oppose singleton and geminate consonants cross-linguistically (see Section 1.2).

(5) *Temporal parameters*

    a. CLOSURE DURATION (CD)

        CD was defined as the interval between the offset of the preceding vowel's F2 (i.e. the onset of consonant closure) and the onset of the acoustic release of the stop, or, if no release was identifiable, the onset of the following vowel's F2. CD constitutes the most robust acoustic correlate of gemination cross-linguistically: phonological length systematically corresponds to a longer CD in the world's languages. However, previous work questions this correlation in Somali (Section 1.2). CD of word-internal singleton and geminate /b d g/ will therefore be measured and systematically compared. In addition, the values of CD obtained for word-initial /b d g/ in the three prosodic contexts tested in the experiment will systematically be compared to those obtained for singleton and geminate /b d g/. This comparison will make it possible to determine whether word-initial singleton /b d g/ are as long as geminates on the one hand, and whether their CD gradually increases with the level of the prosodic hierarchy, or is kept constant, on the other hand.

    b. RELEASE DURATION (RD)

        When an identifiable release was observed, RD was defined as the interval between the onset and the offset of the acoustic release burst. As mentioned in Section 1.2, RD is a parameter that opposes singletons and geminates in certain languages, with geminates realized with a longer RD than singletons. In Somali, word-internal singleton /b d g/ have been reported to display a release only sporadically. RD is thus expected to be at best a secondary correlate of gemination. By contrast, word-initial singleton /b d g/ have been reported to be consistently realized with a release. RD may thus prove crucial to characterize a potential phenomenon of domain-initial strengthening. We will therefore determine whether word-initial /b d g/ and lexical /bb dd gg/ display similar values for RD, and whether RD of word-initial /b d g/ increases with the prosodic hierarchy, or remains constant.

    c. DURATION OF THE VOWEL PRECEDING THE TEST STOP (VD)

        VD was defined as the interval between the onset and the offset of the vowel's F2. In some languages, the distinction between singletons and geminates is secondarily marked by a difference in the duration of the preceding vowel. We will therefore consider this parameter, with a view to establishing whether and to which extent it is involved in Somali, in particular in (prosodic) word-internal contexts. If so, is it a secondary correlate of gemination, or does it override CD? The role of VD in a potential domain-initial strengthening process will be less easy to establish, because the vowel preceding the test stop is, by design, associated with different types of prosodic boundaries in the different conditions of the experiment. However, VD can be used to assess whether word-initial /b d g/ are geminated or strengthened in nominal compounds. Compounds, indeed, constitute a single Prosodic Word, on a par with simplex nouns (see Section 1.1.2). The syllabic context of the vowel preceding the test stop is thus identical in both cases: /i/ is located at the end of a word-internal syllable. VD before word-initial /b d g/ of N2 in N1 N2 compounds can therefore be compared with VD before word-internal singleton and geminate /b d g/. By contrast, the vowel preceding the test stop corresponds to the end of a Prosodic Word in genitives, and to the end of a Phonological Phrase in subject–object sequences. In these configurations, this vowel is likely to be subject to PREBOUNDARY LENGTHENING. Prebindary lengthening is a process whereby a segment located immediately before a prosodic boundary n tends to be longer than a segment located before a prosodic boundary $n-1$. This process has been observed in many unrelated languages (see e.g. Grosjean & Deschamps 1972 for French, Wightman et al. 1992 for English, Cho 2011 for a general overview),

and could very well be at work in Somali, too. This means that final /i/ of N1 in N1 N2 genitives and subject–object sequences is likely to be gradually lengthened, independently of the properties of the following consonant. VD thus cannot be used as a clear parameter to assess the status of word-initial /b d g/ of N2 in N1 N2 genitives and subject–object sequences.

(6) *Non-temporal parameters*

    a. Presence/absence of release
    This parameter notes whether a stop is produced with or without a burst. This will be established through visual inspection of the spectrograms. In Somali, like in other languages, geminates and word-initial /b d g/ have been reported to be systematically realized as true stops with a clear release burst, whereas such a release is reported to be absent, or very weakly realized, in word-internal /b d g/. We will examine this parameter with a view to establishing to which extent it opposes geminate and word-initial /b d g/ on the one hand vs. word-internal /b d g/ on the other hand.

    b. (De)voicing of the test stop
    Voicing is defined as the presence of a voicing bar on the spectrogram. This will be established through visual inspection of the spectrograms. A given consonant will be considered fully voiced if voicing is not interrupted during the production of the consonant, i.e. if the voicing bar is present without interruption between the offset of the preceding vowel and the onset of the following vowel. (This includes the closure and the potential release burst of the consonant.) A consonant will be considered partially (de)voiced if voicing is interrupted during the interval of the consonant. Following standard assumptions (e.g. Ladefoged & Maddieson 1995), we will consider a given consonant to be partially voiced if voicing occurs during the LAST part of the closure only (i.e. if voicing starts before or at the same time as the release), and to be partially devoiced if voicing occurs during the FIRST part of the closure and ends before the release. A consonant will be analysed as fully devoiced if there is no visible voicing bar during the interval of the consonant. As seen in Section 1.1, partial devoicing goes along with gemination in several languages, and must therefore be systematically examined.

    c. Stop closure amplitude (CA)
    The Root Mean Square (RMS) amplitude will be measured over the closure duration of singleton and geminate /b d g/. Amplitude is not an absolute value: it may be affected by various factors, including the general loudness of the utterance for instance. To allow for comparisons across utterances, the RMS values will be normalized by dividing the RMS measured around the center of the closure portion (center + and –20% of the total time of the closure) by the RMS of the center of the preceding vowel (+ and –20% of the total duration of the vowel) and multiplied by 100. Stop closure amplitude is used to assess the degree of constriction of the consonant: the greater the constriction, the smaller the RMS amplitude. The stop closure amplitude of geminates is thus expected to be smaller than that of singletons. In addition, if word-internal singleton /b d g/ are realized as approximants, they are expected to display the highest values for this parameter. Indeed, approximants are characterized by a higher level of energy because of their larger degree of openness (see for instance Martínez-Celdrán & Regueira 2008 and Figueroa Candia 2016 for this effect in different varieties of Spanish). Finally, we will determine the position of word-initial /b d g/ with respect to CA in the spectrum ranging from word-internal singletons to geminates.

    d. Stop release amplitude (RA)
    When an identifiable release was observed, the RMS values were normalized by dividing the RMS of the release portion by the RMS of the center of the following vowel
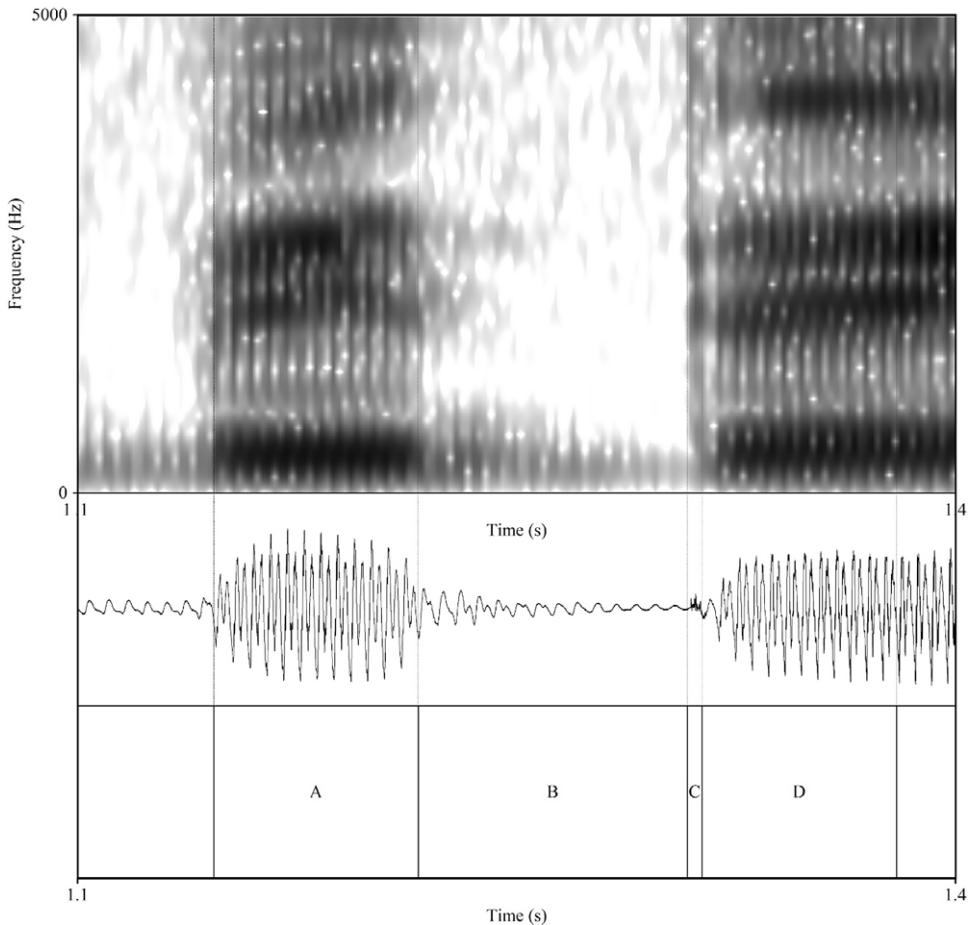
Figure 1 Segmentation and labelling of the sequence /bidda/ in *biddayaasha* 'the male slaves'.

(+ and –20%) and multiplied by 100. RA will be examined on the same grounds that lead us to examine RD: like RD, RA may be involved in the contrast singleton vs. geminate, with geminates realized with a greater RA than singletons. A comparison of the values obtained for word-initial /b d g/ with those of /bb dd gg/, as well as whether these values are correlated with the prosodic hierarchy will be relevant for the characterization of the phenomenon at work word-initially (gemination, gradual strengthening, or neither).

The segmentation procedure is exemplified in Figure 1 for the sequence /bidda/ in *biddayaasha* 'the male slaves'. A corresponds to the vowel /i/ preceding the test consonant, B corresponds to the closure of the test consonant /dd/, C corresponds to its release burst and D corresponds to the vowel /a/ following the test consonant.

2.4   Data taken into account and statistical analysis

A close examination of the recordings led us to exclude four main groups of data that deserve being treated separately. First, the corpus was designed so as to exclude word prominence on the vowel preceding the test consonant. However, some speakers produced patterns that diverge from the standard assumptions on the distribution of Somali tonal accent (see

Table 9 Number of items taken into account in the statistical analysis (in black) for each context and consonant. GenCmp (in grey) was excluded.

Context		b	d	g	Total
LexC	(word-internal singleton stop)	74	71	72	217
LexCC	(word-internal geminate stop)	50	62	60	172
Cmp	(initial singleton stop of N2 in N1 N2 compounds)	53	64	73	190
Gen	(initial singleton stop of N2 in N1 N2 genitives)	51	60	52	163
Ind	(initial singleton stop of N2 in N1 N2 subject-object sequences)	11	16	18	45
GenCmp	(initial singleton stop of N2 in N1 N2 genitives realized as compounds)	8	20	32	60
Total		239 (247)	273 (293)	275 (307)	787 **(847)**

the references given in Section 1.1.2), in particular in subject–object and genitive N1 N2 sequences.[16] We excluded these sentences with unexpected prominence on the vowel preceding the test consonant. The second group includes sentences with a pause before the test consonant. In this context, the identification of the closure onset of the test stop is problematic (Farnetani 1981, Flege 1982, Solé 2018). Such is particularly the case in subject–object N1 N2 sequences. The third group includes sentences with fuzzy boundaries between the test stop and the surrounding vowels: this situation mostly arises in the case of word-internal singleton consonants. In this configuration it was impossible to clearly identify the closure of the test stop. Finally, the fourth group includes indefinite genitive constructions that were consistently produced with a single word prominence, i.e. were realized as compounds. We consider this subset, labelled GenCmp in Table 9, to constitute a specific condition whose status is not immediately clear. We will not take it into account, and leave a comparison of GenCmp with the other conditions for further research. As a result, 787 items were taken into account in the statistical analysis. The number of items for each consonant in each condition appears in Table 9.

Two non-temporal parameters (presence/absence of release and devoicing) were descriptively analysed calculating their relative frequency, for each context and consonant. The other parameters were studied via a statistical analysis using Linear Mixed-effects Models (henceforth LMM), which provide a powerful tool for the analysis of grouped data (Baayen, Davidson & Bates 2008, Cunnings 2012, among others). LMMs were performed using R (R Core Team 2019) with the packages *lme4* (Bates et al. 2015) and *lmerTest* (Kuznetsova, Brockhoff & Christensen 2017), which provide *p*-values in type I, II, or III ANOVA and summary tables for *lmer* model fits via Satterthwaite's degrees of freedom method. The *Context* (Cmp, Gen, Ind, LexC and LexCC) and the *Consonant* type (b, d, g) were included as fixed factors predicting the measured parameters: three temporal parameters (closure, release and vowel duration), and two non-temporal parameters (closure and release amplitude). As random effects, intercepts for Speakers and Item Repetitions were modelled. The REML (Restricted Maximum Likelihood) method was applied, and once the model was estimated, it was adjusted with the ML (Maximum Likelihood) method, using the update function of *lmerTest* (Cauquil & Combes 2019).

---

[16] As mentioned in Section 1.1.2, the distribution of Somali word prominence and boundary H tone has been understudied until now. In particular, dialectal variation has not been taken into account. Our data underline the fact that more studies are needed in this area.

3 Results

3.1 Temporal parameters

3.1.1 Closure duration (CD)

The mean CD values appear as descriptive plots in Figure 2. (See also Table A1 in the appendix.)

As expected, word-medial singleton consonants (LexC) have a shorter CD than word-medial geminates (LexCC): geminates are twice as long as their singleton counterparts for /b/ (106%) and around 70% longer for /d/ and /g/. CD has comparable values in the three word-initial contexts (Cmp, Gen and Ind). In particular, none of these three conditions exhibits a marked increase of CD. In addition, the value of CD in word-initial position is similar to that of lexical geminates. Further examination of the data indicates that /g/ seems to be a little shorter than /b/ and /d/ in all conditions except in the word-medial singleton context (LexC), and that /b/ is longer when it is geminated.

The statistical analysis confirms these observations: there is a significant effect of the *Context* ($\chi^2(4) = 385.96$, $p < .001$), of the *Consonant* ($\chi^2(2) = 14.47$, $p < .001$) and a significant *Context∗Consonant* interaction ($\chi^2(8) = 39.91$, $p < .001$). Pairwise comparisons (Tukey) for the *Context* effect show a significant difference between the word-medial singleton context (LexC) on the one hand, and all other conditions on the other hand (relative to Cmp: $\beta = 35.45$, $se = 1.46$, $t = 24.21$, $p < .0001$; to Gen: $\beta = 33.30$, $se = 1.52$, $t = 21.93$, $p < .0001$; to Ind: $\beta = 34.63$, $se = 2.46$, $t = 14.07$, $p < .0001$; to LexCC: $\beta = 35.47$, $se = 1.50$, $t = 23.57$, $p < .0001$). There is no significant difference between the word-medial geminate context (LexCC), the N1 N2 compound context (Cmp), the N1 N2 genitive context (Gen) and the N1 N2 subject–object context (Ind). The significant effect of the Context is thus to be ascribed to the word-medial singleton context (LexC) alone.[17] Pairwise comparisons (Tukey) for the main *Consonant* effect show a significant difference between /g/ vs. /b/ ($\beta = 6.86$,
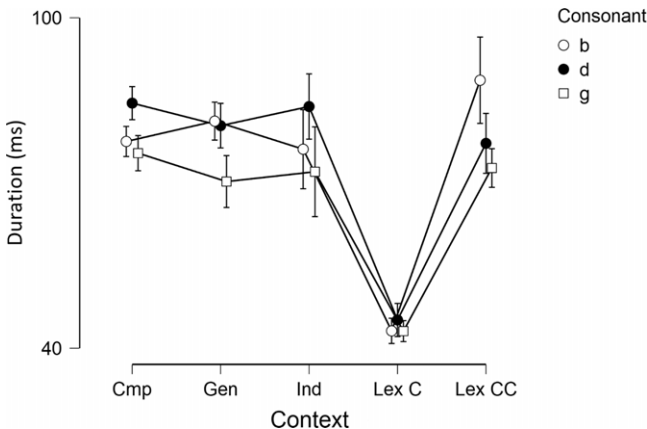

Figure 2 Plots for mean CD (ms) with confidence interval (95%).

---

[17] The same generalization holds if each consonant is taken separately: for the significant *Context∗Consonant* interaction, we report a significant effect of the context for each of the three stops (/b/: $\chi^2(4) = 355.98$, $p < .001$; /d/: $\chi^2(4) = 309.70$, $p < .001$; /g/: $\chi^2(4) = 276.49$, $p < .001$). Post-hoc pairwise comparisons (Tukey) show that, for each consonant, there is a significant difference between LexC on the one hand, and all other conditions on the other hand ($p < .0001$). For each consonant, LexCC, Cmp, Gen and Ind do not differ in any significant way, except for one single case: /b/. In this case, the interaction is due to a significant difference between LexCC and Cmp ($\beta = -10.95$, $se = 3.00$, $t = -3.647$, $p < .01$).

$se = 1.56$, $t = 4.40$, $p < .0001$) and /d/ ($\beta = 7.72$, $se = 1.45$, $t = 5.32$, $p < .0001$). /b/ and /d/ do not differ in any significant way. This result can be ascribed to the aerodynamic configuration of /g/. Voicing requires the subglottal pressure to exceed the intraoral pressure by a threshold value (see e.g. Titze 1988). The intraoral pressure crucially depends on the volume and net compliance of the surfaces above the glottis (e.g. Ohala & Riordan 1979, Ohala 1983,[18] Solé 2018). During the production of /g/, less supra-glottalic volume and soft surfaces are available when compared to those of /d/ and /b/; oral pressure thus increases more quickly until it reaches the critical voicing threshold when release obtains. Accordingly, the closure is released earlier for /g/ than for /b/ and /d/ in order to maintain voicing.

To sum up, the statistical analysis establishes three main results: (i) CD of word-internal singleton voiced stops systematically differs from that of geminates and word-initial singleton voiced stops: it is significantly shorter, and this is true for all consonants; (ii) CD of word-initial stops is comparable to that of geminates, and (iii) CD of word-initial singleton voiced stops has the same value in various prosodic contexts: no significant lengthening – or shortening – has been observed at the beginning of Noun2 in any particular context.

3.1.2 Release duration (RD)

The mean RD values appear as descriptive plots in Figure 3. (See also Table A1 in the appendix.)

The first observation concerns the word-medial singleton context (LexC): only 6% of all test stops ($N = 13/217$) were realized with a release burst in this condition. For this reason, it was not taken into account in the statistical analysis of RD. The second observation is that in all conditions, RD of /b/ is shorter than that of /d/, which is shorter than that of /g/. There seems to be no effect of the context for /b/ and /d/. By contrast, there seems to be an effect of the context for /g/: shorter RD in the N1 N2 subject–object context (Ind) and in the word-medial geminate context (LexCC). However, since RD is characterized by a great variation for /g/, this observation should be taken cautiously.
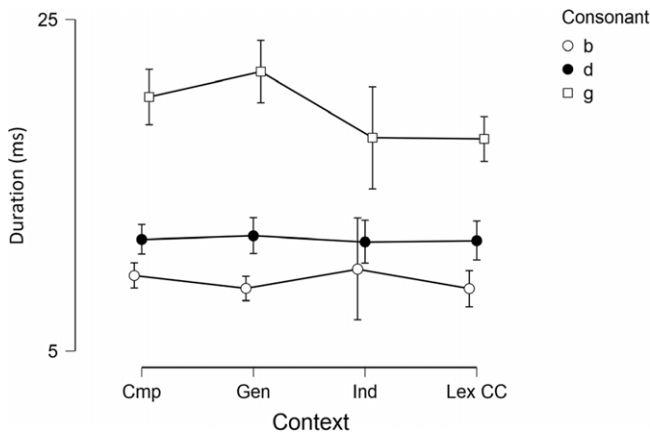

Figure 3 Plots for mean RD (ms) with confidence interval (95%).

---

[18] Ohala (1983: 197): '[the duration in voicing] can be accounted for by considering the net compliance of the surfaces on which oral air pressure impinges during the production of the stops. For velar stops only the pharyngeal walls and part of the soft palate can yield to the air pressure; in dentals, these surfaces plus the greater part of the tongue surface and all of the soft palate are involved; and in labials, these surfaces plus all of the tongue surface and some parts of the cheeks participate'.

These observations are supported by the statistical analysis, which establishes a significant main effect of *Consonant* ($\chi^2(2) = 191.44$, $p < .001$). Pairwise comparisons (Tukey) show significant differences between all consonants: b vs. d ($\beta = -2.40$, $se = 0.617$, $t = -3.887$, $p < .001$); b vs. g ($\beta = -10.11$, $se = 0.623$, $t = -16.228$, $p < .0001$); and d vs. g ($\beta = -7.72$, $se = 0.555$, $t = -13.893$, $p < .0001$). This *Consonant* effect, even if significant, is not important in absolute value. The difference between /b/, /d/ and /g/ probably reflects a simple articulatory effect, that does not trigger the perception of consonant lengthening. In addition, we report no main effect of the *Context*. In particular, word-internal geminates and word-initial stops have similar RD.[19]

In sum, two results obtain: (i) word-internal singleton /b d g/ are realized with no release, and (ii) RD does not seem to be a parameter that clearly distinguishes word-initial singleton stops and word-internal geminates. As was the case with CD, RD is independent from the context: for a given consonant, it remains constant in all contexts.

3.1.3 Vowel duration (VD)

The mean VD values appear as descriptive plots in Figure 4. (See also Table A1 in the appendix.)

We report a significant *Context* effect ($\chi^2(4) = 56.32$, $p < .001$). Follow-up pairwise comparisons (Tukey) reveal no significant difference between the word-medial singleton context (LexC), the word-medial geminate context (LexCC) and the N1 N2 compound context (Cmp), but significant differences between these conditions and the N1 N2 subject–object context (Ind) and the N1 N2 genitive context (Gen): Cmp vs. Gen ($\beta = -15.04$, $se = 1.79$, $t = -8.382$, $p < .0001$), Cmp vs. Ind ($\beta = -23.31$, $se = 2.68$, $t = -8.697$, $p < .0001$), Gen vs. LexC ($\beta = 12.97$, $se = 1.82$, $t = 7.126$, $p < .0001$), Gen vs. LexCC ($\beta = 13.24$, $se = 1.99$,
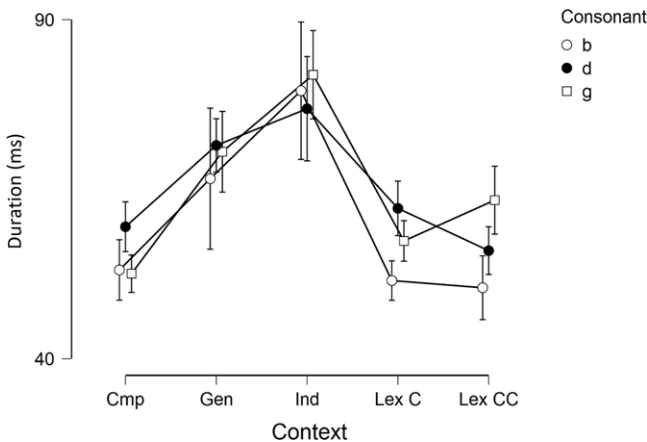

Figure 4 Plots for mean VD (ms) with confidence interval (95%).

---

[19] There are no significant differences for /b/ and /d/. For /g/, there are no significant differences between Cmp, Ind and LexCC (Cmp and Gen ($\beta = -1.411$, $se = 1.22$, $t = -1.157$, $p = .6546$); Cmp and Ind ($\beta = 3.140$, $se = 1.72$, $t = 1.829$, $p = .2629$); Cmp and LexCC ($\beta = 2.437$, $se = 1.18$, $t = 2.065$, $p = .1681$); Ind and LexCC ($\beta = -0.703$, $se = 1.85$, $t = -0.380$, $p = .9813$), but significant differences between Gen and LexCC ($\beta = 3.848$, $se = 1.35$, $t = 2.846$, $p < .05$) and Gen and Ind ($\beta = 4.551$, $se = 1.71$, $t = 2.657$, $p = .0421$). These differences (i) only concern /g/, and (ii) consist in variations of 3–4 ms, which are too short to be under the control of the speakers and should have no perceptual effect. Finally, we report a significant *Context∗Consonant* interaction ($\chi^2(8) = 16.94$, $p < .01$), which can be ascribed to the particular pattern of /g/.

$t = 6.663$, $p < .0001$), Ind vs. LexC ($\beta = 21.25$, $se = 2.70$, $t = 7.862$, $p < .0001$), and Ind vs. LexCC ($\beta = 21.51$, $se = 2.82$, $t = 7.632$, $p < .0001$). VD of Gen and Ind significantly differ ($\beta = -8.278$, $se = 2.72$, $t = -3.042$, $p < .05$). In sum, we reach the following pattern for VD: word-medial singleton (LexC) $\approx$ word-medial geminate (LexCC) $\approx$ N1 N2 compound (Cmp) < N1 N2 genitive (Gen) < N1 N2 subject object (Ind).[20]

We are now in a position to establish three important generalizations. (i) There is no significant difference in VD before word-medial singleton and geminate stops: VD is not a factor that opposes singleton and geminate stops in Somali. In this sense, Somali behaves like Turkish for instance. (ii) VD before the initial stop of N2 in N1 N2 compounds is comparable to VD before a word-internal singleton or geminate stop. (iii) VD increases with the hierarchical level of the prosodic boundary located before N2: the vowel located immediately before a Phonological Phrase boundary in N1 N2 subject–object sequences is longer than the vowel located before a Prosodic Word boundary in N1 N2 genitive constructions. Our results thus clearly establish that Somali displays the well-attested phenomenon of pre-boundary lengthening: a vowel located immediately before a prosodic boundary n is longer than a vowel located immediately before a prosodic boundary $n-1$ (Section 2.3).

3.2 Non-temporal parameters

3.2.1 Stop closure amplitude (CA)

The mean CA values are given in Figure 5 as descriptive plots. (See also Table A2 in the appendix.)

The data are characterized by a great variation, in particular in the N1 N2 subject–object context (Ind), in the word-medial singleton context (LexC) and in the word-medial geminate context (LexCC). We report no significant effect of *Consonant* ($\chi^2(2) = .871$, $p = .647$). By contrast, there is a main *Context* effect ($\chi^2(4) = 218.64$, $p < .001$). Pairwise comparisons (Tukey) show a significant difference between the word-medial singleton context (LexC) and
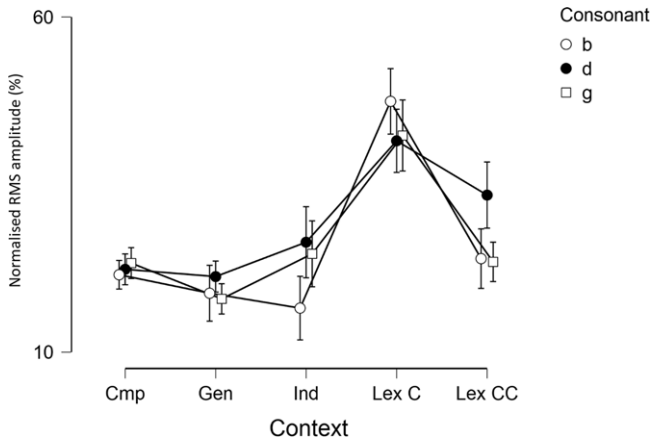

Figure 5 Plots for mean CA (%) with confidence interval (95%).

---

[20] There are only marginal effects for *Consonant* ($\chi^2(2) = 6.719$, $p < .05$) and the *Context∗Consonant* interaction ($\chi^2(8) = 16.86$, $p < .05$). For the *Consonant* effect, pairwise comparisons show significant differences between /b/ and /d/ ($\beta = -5.19$, $se = 1.76$, $t = -2.954$, $p < .01$) and /b/ and /g/ ($\beta = -5.15$, $se = 1.78$, $t = -2.890$, $p < .05$) and no difference between /d/ and /g/. As for the *Context∗Consonant* interaction, it is to be ascribed to the fact that each consonant has a specific VD pattern in word-internal contexts.

all other conditions (relative to LexCC: $\beta = -16.64$, $se = 1.37$, $t = -12.11$, $p < .0001$; to Cmp: $\beta = -21.48$, $se = 1.33$, $t = -16.14$, $p < .0001$; to Gen: $\beta = -24.91$, $se = 1.38$, $t = -17.99$, $p < .0001$; to Ind: $\beta = -25.33$, $se = 2.24$, $t = -11.30$, $p < .0001$). In addition, there are significant differences between the word-medial geminate context (LexCC) and all word-initial conditions (relative to Cmp: $\beta = -4.94$, $se = 1.41$, $t = -3.50$, $p < .001$; to Gen: $\beta = -8.361$, $se = 1.46$, $t = -5.71$, $p < .0001$; to Ind: $\beta = -8.79$, $se = 2.30$, $t = -3.83$, $p < .01$). However pairwise comparisons (Tukey) of the contexts for each consonant show that the word-medial geminate context (LexCC) differs from the word-initial contexts for /d/, only (relative to Cmp: $\beta = -10.93$, $se = 2.48$, $t = -4.41$, $p < .0001$; to Gen: $\beta = -12.77$, $se = 2.52$, $t = -5.06$, $p < .0001$; to Ind: $\beta = -11.50$, $se = 3.98$, $t = -2.90$, $p < .05$).[21] Finally, no significant difference obtains between the N1 N2 subject–object context (Ind), the N1 N2 genitive context (Gen) and the N1 N2 compound context (Cmp).

To sum up: (i) word-internal singleton stops exhibit the highest ratio of CA (41.5–47.4%); they clearly differ from geminates (23.5–33.4%) and word-initial singleton stops (16.6–26.4%); (ii) CA of word-initial singleton stops is comparable in all three contexts; (iii) abstracting away from the peculiarity of /d/, CA of geminates and CA of word-initial singleton stops are comparable. In this sense the results are the mirror image of those obtained for CD: the shorter CD is, the higher the amplitude of the closure is.

3.2.2 Release amplitude (RA)

Figure 6 (see also Table A2) displays the mean RA values in the word-initial contexts and the word-medial geminate context (LexCC).[22]

The results for RA are characterized by a great variability: Figure 6 reveals high values for confidence intervals, see also high values for SD in Table A2. Consider first /d/ and /g/. RA of /g/ is consistently lower than that of /d/. RA of /g/ and /d/ vary in the same way with respect to the context: RA is proportional to the strength of the prosodic boundary before N2 (the stronger the boundary is, the higher RA is). This suggests an influence of the prosodic structure on the realization of voiced stops, and more specifically a phenomenon
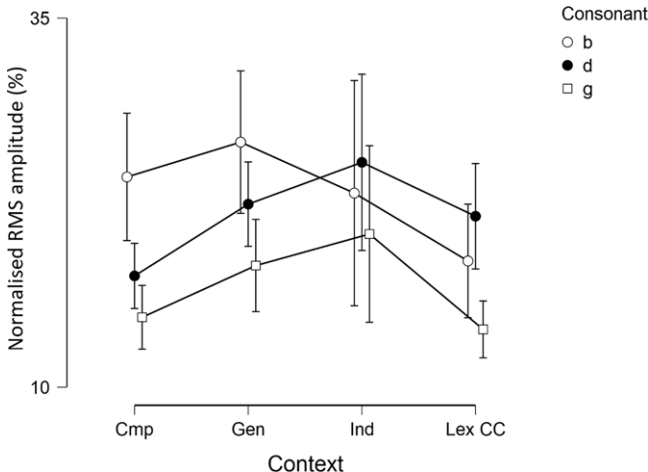

Figure 6 Plots for mean RA (%) with confidence interval (95%).

---

[21] For /b/ and /g/, LexCC differ from the word-initial contexts in a single case, and with $p < .05$ only: LexCC vs. Gen for /g/ ($\beta = -6.54$, $se = 2.23$, $t = -2.93$). Finally, a significant *Context∗Consonant* interaction is observed ($\chi^2(8) = 28.41$, $p < .001$), which can be ascribed to the peculiar behaviour of /d/.
[22] LexC was excluded because there was no burst in this condition (Section 3.1.2).

of domain-initial strengthening that would be cumulatively implemented via RA. This conjecture however must be tempered by two additional observations. First, /g/ and /d/ do not behave exactly in a parallel way: for /g/, the word-medial geminate context (LexCC) is similar to the N1 N2 compound context (Cmp), but for /d/ it is similar to the N1 N2 genitive context (Gen). Second, the putative domain-initial strengthening phenomenon observed for /d/ and /g/ does not seem to obtain with /b/: there, the values obtained for the N1 N2 subject–object context (Ind) are lower than those obtained for the N1 N2 compound and genitive contexts (Cmp and Gen resp.), and the variability of the results is even higher. We thus need a statistical analysis to confirm this point.

The statistical analysis shows a significant effect of the *Consonant* ($\chi^2(2) = 28.67$, $p <$ .001), of the *Context* ($\chi^2(3) = 19.41$, $p < .001$), and a significant *Context*∗*Consonant* interaction ($\chi^2(6) = 13.46$, $p < .05$). For the *Consonant* effect, post-hoc tests (Tukey) show that /g/ systematically differs from /b/ and /d/: there are no significant differences between /b/ and /d/, but significant differences between /g/ and /b/ ($\beta = 6.32$, $se = 1.28$, $t = 4.95$, $p < .0001$) and between /g/ and /d/ ($\beta = 4.70$, $se = 1.14$, $t = 4.12$, $p < .0001$). This is coherent with what has been observed for RD, and is to be ascribed to physiological and aerodynamic conditions: the total amount of air present in the oral cavity is less important for /g/ than for /b/ and /d/; thus, when the articulators separate from each other, less air is released and the release burst is weaker. As for the *Context* effect, the word-initial contexts (Gen, Ind and Cmp) do not significantly differ from each other. Moreover, RA does not provide a clear-cut pattern for the word-medial geminate context (LexCC).[23] This weakens the hypothesis of a domain-initial strengthening phenomenon, which is expected to be correlated with the prosodic hierarchy.

To conclude, RA reveals an intricate pattern. In particular, if the observation of the data points to a phenomenon of domain-initial strengthening, the statistical analysis does not confirm this effect.

Interestingly, our results indicate a contrast between, on the one hand, the temporal parameters and, on the other hand, the amplitude parameters: the word-medial geminate context (LexCC) and the word-initial contexts (Ind, Cmp and Gen) clearly pattern together with respect to the temporal parameters; they less clearly do so with respect to the amplitude parameters.

We now turn to the last two parameters – presence/absence of release and (de)voicing – which were analysed via visual inspection.

3.2.3 Presence/absence of release

A visual inspection of the oscillographic signals and spectrograms shows that 94 % ($N = 204/217$) of the word-medial singletons (LexC) are produced WITHOUT any release burst (for an example, see Figure 7); this holds true for all places of articulation.

By contrast, word-initial singleton stops (Cmp, Gen and Ind) and word-medial geminates (LexCC) are realized with a clearly identifiable release ($N = 567/570$, 99.5 %): see Figures 8, 9 and 10.

---

[23] Pairwise comparisons (Tukey) for the *Context* effect show significant differences between LexCC and Gen, only ($\beta = 5.09$, $se = 1.14$, $t = 4.48$, $p < .0001$), and a marginal significant effect between LexCC and Ind ($\beta = 5.13$, $se = 1.78$, $t = 2.88$, $p < .05$). Pairwise comparisons (Tukey) for each consonant show small significant differences for /b/ and /g/, only (for /b/, LexCC vs. Gen ($\beta = 8.96$, $se = 2.61$, $t = 3.43$, $p < .01$) and LexCC vs. Cmp ($\beta = 7.49$, $se = 2.62$, $t = 2.86$, $p < .05$); for /g/, LexCC vs. Gen ($\beta = 5.09$, $se = 1.75$, $t = 2.92$, $p < .05$) and LexCC vs. Ind ($\beta = 7.04$, $se = 2.49$, $t = 2.83$, $p < .05$). There are no significant differences for /d/. LexCC tends to differ from the word-initial contexts in that it is characterized by a weaker RA. This is not a solid observation, though: for /g/ LexCC does not significantly differ from Cmp; for /b/, LexCC does not significantly differ from Ind; for /d/ LexCC does not significantly differ from any word-initial context.
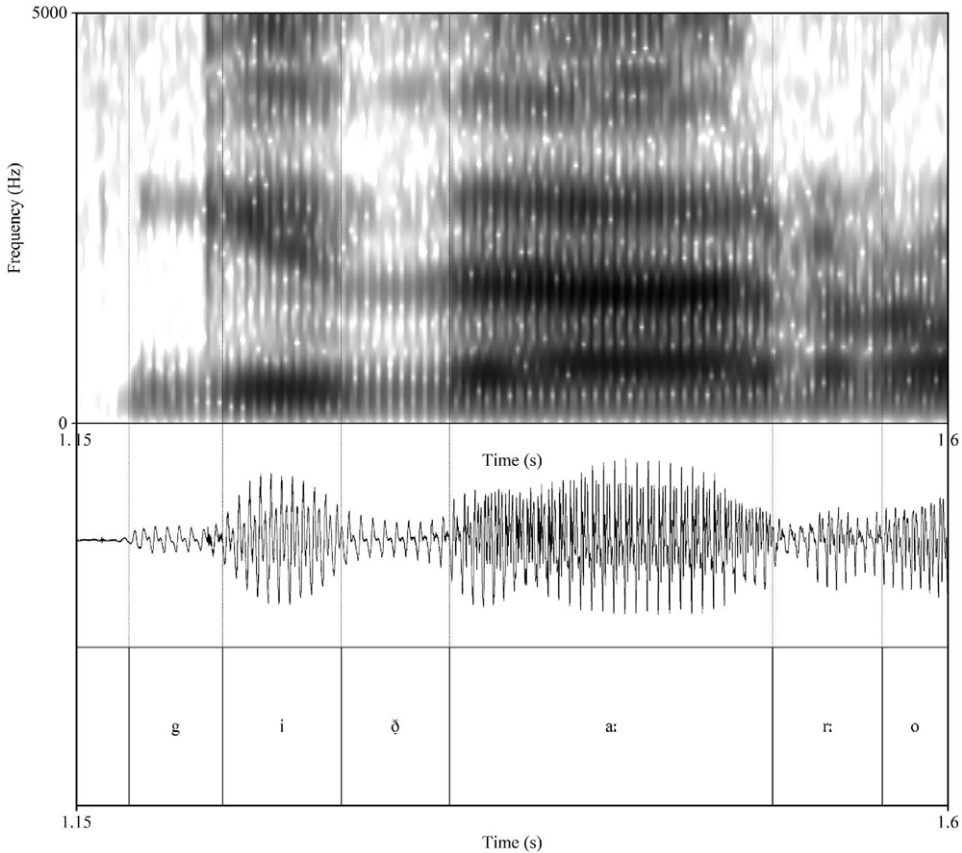
Figure 7 Spectrogram and waveform of *gidaarro* 'walls' showing the realization of /d/ in LexC. /d/ is produced as a voiced approximant (no release burst, formant structure, see Section 4.1.1): [giḏaːrːo].

3.2.4 Devoicing

Devoicing of geminate consonants has been reported for various languages. As mentioned earlier (Section 3.1.1), this phenomenon can be explained by aerodynamic considerations: the longer the CD is, the more air pressure builds up. As a consequence, the transglottal air pressure differential drops below the threshold for voicing. Devoicing during the production of a geminate stop is thus a consequence of its longer CD.

In the word-medial geminate context (LexCC), our data show either a constant level of voicing, or a slight attenuation of voicing. Devoicing is absent, and partial devoicing is marginal (four tokens realized by the same speaker). Since the results obtained for CD suggest that word-initial singleton stops (Cmp, Gen and Ind) behave like word-medial geminates (LexCC), these conditions must be inspected as well. It turns out that, again, word-initial singleton stops pattern with word-medial geminates: devoicing is produced in 23 occurrences of word-initial stops, only (Cmp = 15, Gen = 8, Ind = 0), and by a single speaker. For all other speakers, the stops recorded word-initially are fully voiced, with voicing attenuation in certain cases (see Figure 8 and Figure 9 for illustration).

We thus report a nearly complete absence of devoicing in Somali. By contrast, Farnetani (1981) reports partial devoicing of geminate voiced stops (Section 1.2). There is a straightforward explanation for this discrepancy. The results presented in Section 3.1.1 reveal that CD of geminate stops is strikingly short in our data. In Farnetani's data, CD is much longer. It thus just seems to be the case that the geminates that we recorded have a CD that is too short to create an intraoral pressure that is sufficiently high to trigger devoicing. We conclude
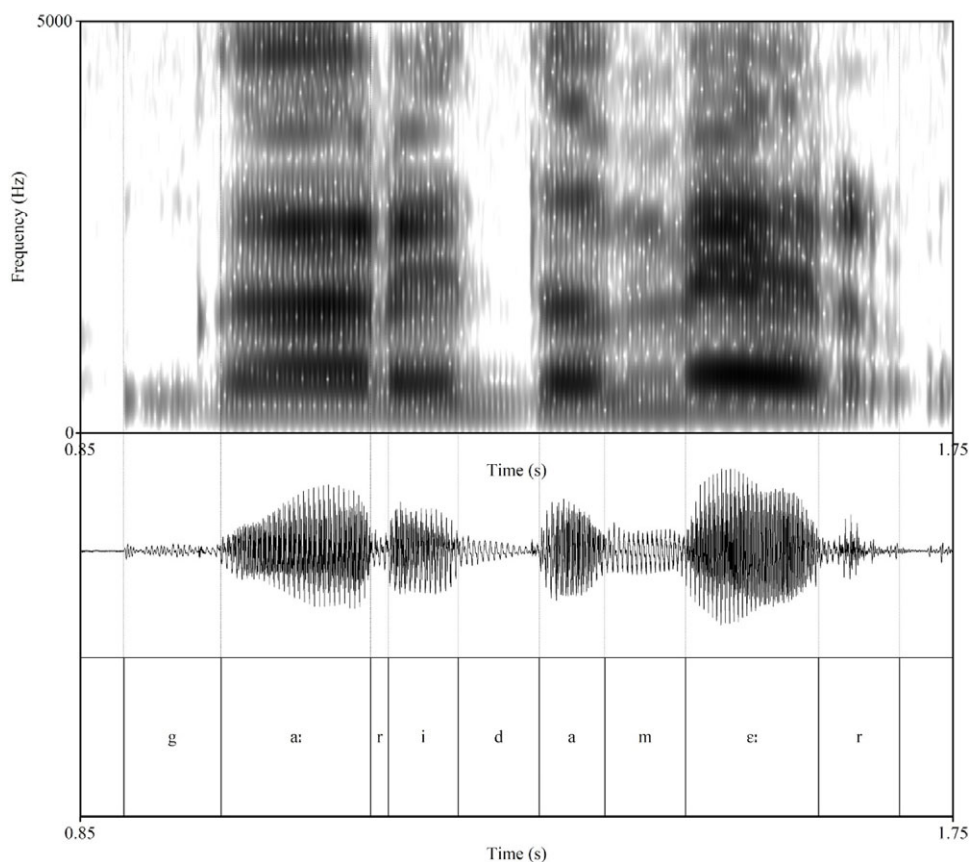
Figure 8 Spectrogram and waveform of *gaari-dameer* 'a cart pulled by a donkey' showing the realization of /d/ in Cmp. /d/ is produced as a voiced stop with a release burst: [gaːridamɛːr].

that devoicing of geminate voiced stops is directly correlated to CD in Somali. This has two implications: (i) this phenomenon cannot be analysed as a change in the phonological specification for voicing of the stop in certain contexts, and (ii) the variation observed in the literature on Somali is to be ascribed to a variation in CD.

4 Discussion

In this section, we first examine the implication of our results on the contrast between singletons and geminates (Section 4.1), and then we address the question of the word-initial position (Section 4.2). For each issue, we discuss the results obtained in the acoustic analysis, and evaluate their implications at the phonological level. Since the present study is based on a small number of speakers, the discussion in this section is meant as a preliminary analysis and constitutes the basis for further research.

4.1 Word-medial singleton vs. geminate voiced stops

4.1.1 The acoustic correlates of singleton vs. geminate voiced stops

Somali word-medial singleton /b d g/ are characterized by a short duration (43.1–45.1 ms), the absence of a release burst, and a high closure amplitude. High closure amplitude
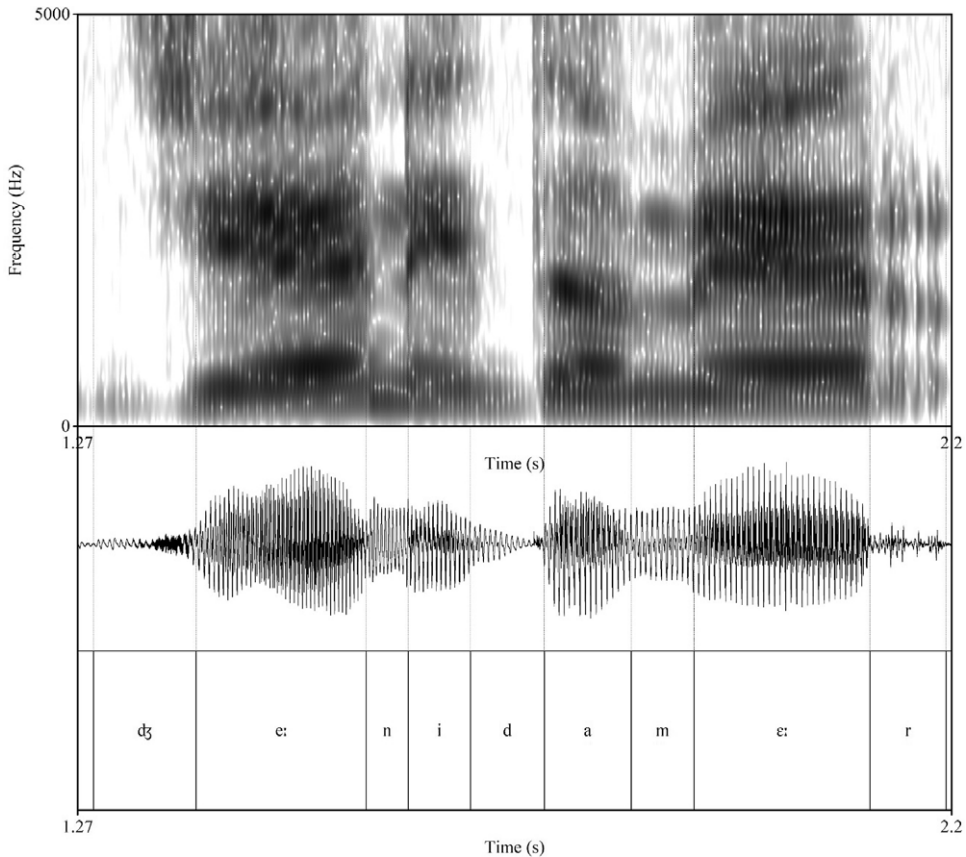
Figure 9 Spectrogram and waveform of *jeeni dameer* 'front leg of a donkey' showing the realization of /d/ in Gen. /d/ is produced as a voiced stop with a release burst: [d͡ʒeːnidameːr].

corresponds to a high degree of articulatory openness. This characteristic can also be established via visual inspection of the spectrograms: word-internal /b d ɡ/ display formant structures and a high level of energy (see Figure 7). We can safely conclude that word-internal /b d ɡ/ are realized as approximants: [β ð̞ ɣ]. Comparable results have been obtained in various unrelated languages, e.g. different varieties of Spanish.[24] Martínez-Celdrán & Regueira (2008) distinguish three subclasses in the approximant category: closed, open and vocalic approximants. These subclasses are distinguished on the basis of acoustic characteristics that correspond to three degrees of articulatory openness. Adopting this terminology, we can state that Somali word-internal /b d ɡ/ exhibit the acoustic properties of open approximants: they are short and their formant structure represents a transition between the adjacent vowels with clear glottal pulses above the voice bar, see Figure 7 above for an example. In addition, recall that several test consonants could not be segmented because the boundaries with the adjacent vowels were too fuzzy (Section 2.4). In these realizations, the articulators merely approach each other and the approximants are characterized by nearly vocalic properties with a very high level of energy. They can be considered as instances of vocalic approximants. Cases of closed approximants, characterized by weaker glottal pulses, were observed only marginally.

---

[24] For example, Galician Spanish, and Chilean Spanish (Martínez-Celdrán 2004, Martínez-Celdrán & Regueira 2008, Figueroa Candia 2016).
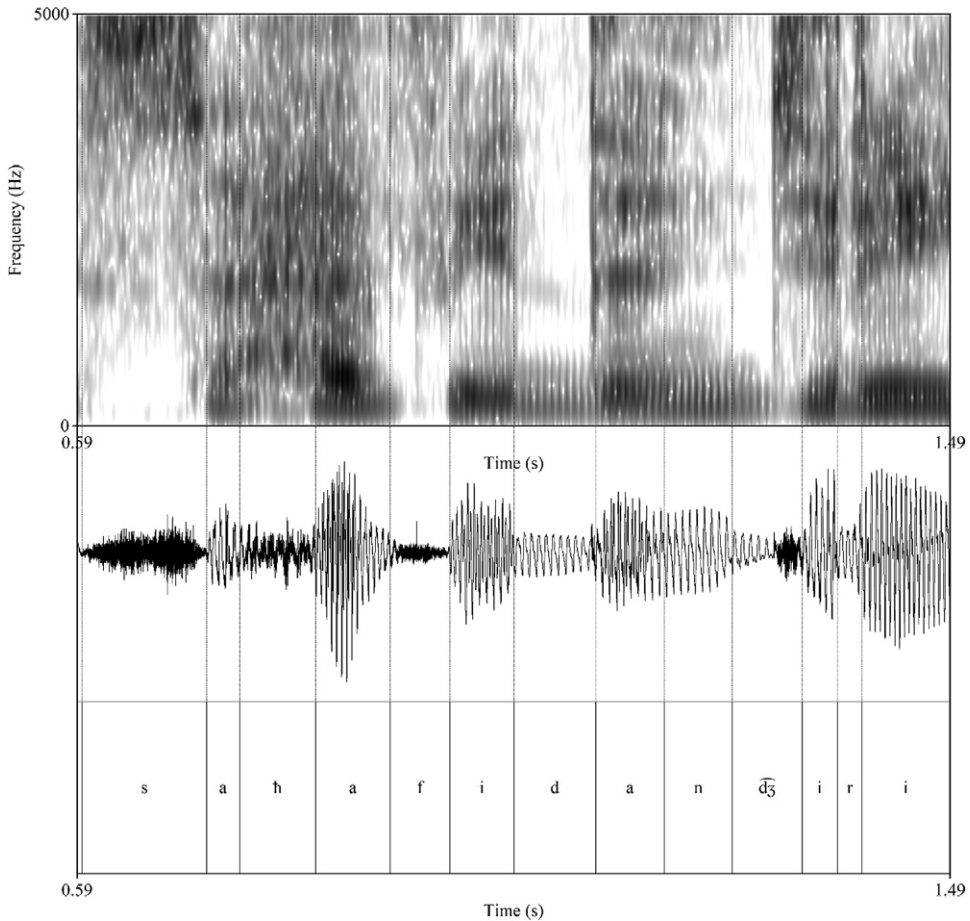
Figure 10 Spectrogram and waveform of *saxafi danjiri*(*hii*) 'journalist ambassador' showing the realization of /d/ in Ind. /d/ is produced as a voiced stop with a release burst: [saħafidandʒiri].

Turning to lexical geminates, /bb dd gg/ are always realized as stops (with a clear release burst), and their closure duration is 70–106% longer than that of their singleton counterparts. For this reason, geminates have a lower closure amplitude than singletons. The other acoustic correlates do not enhance the contrast singleton vs. geminate: Somali does not behave on a par with the languages for which release burst amplitude, devoicing, and/or shortening of the preceding vowel have been reported to distinguish between singletons and geminates.[25] Most strikingly, in our data, the closure duration of geminates is twice as short as that mentioned in Farnetani (1981) (72.7–88.7 ms vs. 157–174 ms respectively). More generally, the values of closure duration obtained for /bb dd gg/ are considerably shorter than those reported for geminates in Afroasiatic languages (e.g. 144 ms in Tashlhiyt Berber, Ridouane 2007), and elsewhere (e.g. 176 ms in Turkish and 255 ms in Bengali, Lahiri & Hankamer 1988). In order to assess the implications of this observation, we reproduce the values reported in Ridouane (2007) for Tashlhiyt Berber in Table 10 and compare them to our values for Somali. (Note that the number of speakers participating in both experiments is comparable:

---

[25] Release duration cannot be considered a parameter in Somali since singleton voiced stops are realized as approximants.

Table 10 Closure duration in milliseconds of singleton and geminate voiced stops in Tashlhiyt Berber (Ridouane 2007: 128) and in Somali.

Language	Initial		Intervocalic	
	Singleton	Geminate	Singleton	Geminate
Tashlhiyt Berber (5 speakers)	75	189	57	144
Somali (4 speakers)	70.3–84.5	DNA	43.1–45.1	72.7–88.7

four in our study and five in that conducted by Ridouane.) In word-initial position, the closure duration of singleton /b d g/ in Somali is comparable to that obtained for /d g dˤ/ in Tashlhiyt Berber. In intervocalic position, Somali singleton and geminate consonants are shorter. This is particularly evident when considering the geminates.

Together with the acoustic properties mentioned earlier, we take this to indicate that lexical geminates in Somali are not realized as geminate stops at the phonetic level, but rather as singleton stops. The contrast word-internal singleton vs. geminate voiced stop is therefore realized as the contrast open approximant vs. singleton stop at the phonetic level.

4.1.2 Phonological implications

This result has implications for the characterization of gemination at the phonetics– phonology interface, and more generally for the question of the isomorphism between phonological representation and phonetic realization. The two following options are in principle available:

(7) a. There is no phonological length contrast in Somali: the phonological system includes two series of consonants characterized by different feature specifications (the approximants and the stops) and no geminates.
   b. There is a phonological length contrast in Somali: phonological length is phonetically realized as a contrast in the manner of articulation, approximant vs. stop.

The predictions made by these two options are summarized in (8) below. An important testing point concerns intervocalic voiced stops, which are phonologically short according to hypothesis (7a), but phonologically long according to hypothesis (7b).

(8) Hypothesis (7a)   Hypothesis (7b)
     　　　V_V   　　　V_V

     [β ð̞ ɣ̞] [b d g]   [β ð̞ ɣ̞] [b d g]

     /β ð̞ ɣ̞/ /b d g/   /b d g/   /bb dd gg/

Various phenomena strongly suggest that intervocalic [b d g] productively behave like CC clusters (see Barillot 2002 and Barillot & Ségéral 2005). Consider for instance the vowel/zero alternations illustrated in (9).

(9) IMPERATIVE.2SG   PRESENT.1SG
                           V/zero alt.   no V/zero alt.
     a. ha[ð̞]al   ha[dˀ]la: *ha[ð̞]ala: 'speak'
     b. be[d]el   *bedla: be[d]ela: 'change'

A verb stem with intervocalic [ð̞], e.g. *hadal* ha[ð̞]al 'speak' in (9a), systematically exhibits vowel/zero alternations in its paradigm. By contrast, a verb stem with intervocalic [d], e.g.

beddel be[d]el 'change' in (9b), never exhibits such vowel–zero alternations. This is due to the fact that CCC clusters are prohibited in Somali: vowel–zero alternations are blocked whenever they would yield a CCC cluster. The absence of alternation in (9b) indicates that [dl] constitutes a CCC cluster, i.e. [d] occupies two consonantal slots. We conclude that intervocalic voiced stops occupy two skeletal slots at the phonological level: they are long, i.e. we validate hypothesis (7b).

The phonological representations of singletons and geminates are given in (10).

(10) a. Singleton   b. Geminate
     x   x x
     | \ /
     b   b
     [β]   [b]

The general question raised here is that of the isomorphism between phonetic realizations and phonological representations. The literature in phonetics tends to support an isomorphism between the number of timing slots assumed at the phonological level and phonetic length: the contrast between geminate and singleton consonants primarily involves a length contrast (see Lahiri & Hankamer 1988 and Ridouane 2010 for a review). At first sight, (10b) implies an absence of isomorphism: there is no length at the phonetic level, but length at the phonological level. However, we would like to offer a more balanced interpretation of (10). First, recall that we report a significant difference in closure duration between word-medial singletons and geminates. The ratio (+70–106%) is comparable to what has been reported for other languages (Ridouane 2010, Hamzah et al. 2016). Second, the closure duration of word-medial singletons remains strikingly stable across places of articulation: it ranges from 43.1 ms to 45.1 ms (Section 3.1.1). Furthermore, additional exploration of the data reveals that closure duration is also extremely stable across speakers (CNA: 40.7 ms, CRX: 41.4 ms, DEE: 46.2 ms, CQA: 46.3 ms). Let us now assume that a skeletal slot (i.e. a timing unit in the phonological representation) corresponds to a given amount of time, which is parametrized. The extremely stable duration of word-medial singletons represents the minimal threshold to produce a consonant. In Somali, a timing unit corresponds to this minimal amount of time. This duration does not provide the articulators with enough time to reach each other, and no occlusion of the vocal tract takes place: word-medial singletons surface as approximants. By contrast, if two timing units are available, then occlusion takes place and a stop surfaces: word-medial geminates are realized as voiced stops. In that sense, the temporal representation of geminates sketched in (10) above does not necessarily imply the absence of isomorphism between phonetics and phonology.

To conclude, contrary to what has been observed in earlier studies, lexical geminates turn out to be strikingly short. With this study, we might witness the transition between the situation of the language nearly 40 years ago, where duration and manner of articulation were relevant, and today, where the manner of articulation seems to be the primary phonetic correlate of gemination. However, as we just argued, the short duration of lexical geminates at the acoustic level is compatible with a temporal representation involving two phonological slots.

4.2 The beginning of the word

4.2.1 The acoustic correlates of word-initial voiced stops

The acoustic properties of word-initial /b d g/ clearly differ from those of word-medial /b d g/: their closure duration is longer and they are realized as voiced stops with a clear release burst in all contexts. Word-initial singleton stops pattern with lexical geminates: same closure

duration, similar release duration, same specification for voice, and a clear release burst. Rather astonishingly, the results obtained for these parameters hold true irrespectively from the prosodic context: word-initial singletons in compounds, genitives and subject–object sequences pattern together with lexical geminates with no significant difference.

Word-initial stops differ from lexical geminates in release amplitude and vowel duration, only. The results obtained for release amplitude, however, display a great variability and are to be taken very cautiously. As for vowel duration, its pattern corresponds to the well-attested phenomenon of preboundary lengthening: it is a direct consequence of the prosodic structures used in the experiment (Section 3.1.3). It is thus independent from the realization of word-initial /b d g/. Note that within the Prosodic Word (compounds), the duration of the vowel preceding word-initial /b d g/ is similar to that of the vowel preceding lexical geminates. This confirms the robust parallelism of word-initial singletons and geminates established above.

4.2.2 Phonological implications

In this section we evaluate the implications of the striking parallelism between word-initial singletons and lexical geminates on the phonological representation of word-initial singletons. The question is whether this parallelism is due to the fact that word-initial singletons are subject to a process of domain-initial strengthening. Adopting for instance the 'Articulatory Undershoot Hypothesis' proposed by Cho & Keating (2001),[26] we could interpret the Somali facts as follows. A word-initial voiced stop is underlyingly short: it is associated to a single skeletal slot (i.e. its phonological representation is the same as the one of a word-internal voiced stop). In order to account for the fact that it is realized as a stop, we must further assume that, because of its word-initial position, the duration of this slot is increased, so that the articulators have time to come in contact with each other, and thus produce a stop. By contrast, in word-internal position, the duration of a skeletal slot is not lengthened, and an approximant surfaces.

However, a closer look at the data reveals that the situation in Somali is more complex. It has been demonstrated in various languages (English, French, Korean, Taiwanese among others), and for the acoustic dimensions that we studied in Somali, that domain-initial strengthening cumulatively increases with the level in the prosodic hierarchy (see Cho 2011 for an overview, Keating et al. 2003, Keating 2006). In Somali, the closure duration and the release duration of word-initial voiced stops are identical in contexts that are embedded in different prosodic structures, *viz.* compounds (N1 Ń2)$_\omega$, genitives ((Ń1)$_\omega$ (Ń2)$_\omega$)$_\phi$, and subject–object sequences ((Ń1)$_\omega$)$_\phi$ ((Ń2)$_\omega$)$_\phi$. This means that the temporal properties of word-initial voiced stops are independent from the type of prosodic boundary they are located at. In other words, there is no gradience in the realization of word-initial voiced stops, a fact that does not seem to be easily reconcilable with the standard properties of domain-initial strengthening. An additional problem comes from the behaviour of nominal compounds. These compounds form a single Prosodic Word. This means that N2 is not preceded by any prosodic boundary. The initial voiced stop of N2 should therefore be realized as a word-medial singleton, i.e. an approximant. However, our results clearly establish that such is not the case.

In sum, Somali word-initial voiced stops have two salient properties: (i) their temporal realization is independent from the prosodic structure, and (ii) they have the same temporal acoustic characteristics as lexical geminates. It seems to be the case that the Prosodic Phonology framework cannot explain these properties without additional stipulations. In this discussion, we would like to offer an analysis that accounts for these properties, makes further

---

[26] A number of proposals have been advanced to account for the articulatory nature of domain-initial strengthening, see Cho (2011: 350) for an overview.

predictions on the grammar of Standard Somali, and relies on phonological tools that have been shown to be involved in various strengthening phenomena in Afroasiatic languages.

In a standard autosegmental phonology framework, temporal properties are encoded at the skeletal level: a skeletal slot encodes a timing unit. The results of our experiment establish that the temporal properties of word-initial singletons are identical with those of lexical geminates. This implies that the left-edge of the noun involves a skeletal unit, which is identified by spreading of the initial stop of the noun. In order to explore this hypothesis, we consider a proposal that has been shown to account for various word-initial strengthening processes, in particular word-initial gemination in Berber verbal morphology (Guerssel 1992) and Biblical Hebrew nominal morphology (Lowenstamm 1996).

In Berber, word-initial gemination is observed e.g. in derived causatives. The causative morpheme is underlyingly a singleton /s-/, as in /faθ/ 'miss', /s-faθ/ → [sfaθ] 'make miss'. However, under specific phonotactic conditions, this prefix surfaces as geminate [sː-], e.g. /xðəm/ 'work, do', /s-xðəm/ → [sːəxðəm] 'make work, do'.

Such a gemination obtains only in word-initial context. Crucially, it is not observed if the causative prefix is preceded by another derivational prefix, e.g. the passive prefix. In order to account for this process, Guerssel (1992) argues that the lexical representation of a derived causative does not only include the skeletal positions that are necessary to accommodate the segmental material of the verb. In addition, it includes an initial empty syllable (O(nset)–N(ucleus)), that assigns the category 'verb' to the structure. This accounts for the gemination of word-initial /s-/ as shown in (11).

(11)

/s-xðəm/ → [sːəxðəm] 'make work, do'
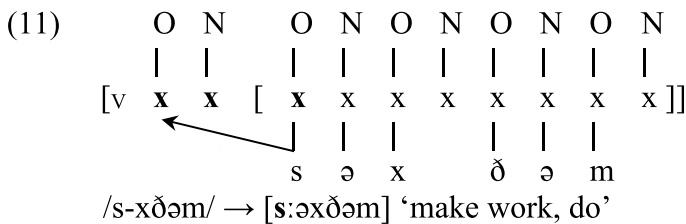
In Biblical Hebrew, when a noun is prefixed by the determiner *ha-* 'the', the initial consonant of the noun must geminate, e.g. /ha-klabim/ 'the-dogs' → [hakːəlabim] 'the dogs'.[27] Lowenstamm (1996) argues that this process results from the spreading of the noun-initial consonant to an initial empty syllable at the beginning of the noun:

(12)

/ha-klabim/ 'the-dogs' → [ha**k**ːəlabim] 'the dogs'
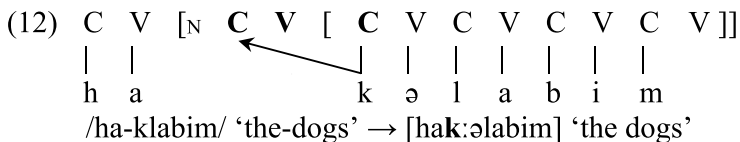
Lowenstamm (1999: 157ff.) further argues that

> Rather than being conventionally marked by the insertion of a # symbol to its left, the word is preceded by an empty CV span. The major difference between this proposal and the traditional view lies in the fact that the initial empty CV span is *a true phonological site*, over which a number of operations will be shown to take place.
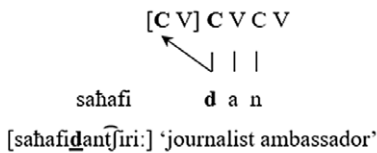
---

[27] If the initial consonant of the noun has no geminate counterpart, the a-vowel of the determiner is lengthened.

According to this proposal, the phonological representation of the major categories, nouns for example, includes an initial skeletal unit.
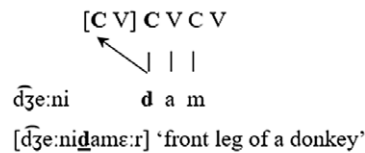
Somali word-initial gemination shares a number of properties with the phenomena illustrated above. In particular, word-initial singletons display the acoustic properties that are characteristics of lexical geminates. It is thus desirable to represent them in the same fashion. Lexical geminates are associated with two skeletal slots. We thus propose to represent word-initial singleton stops as segments associated with two skeletal slots at the phonological level. The proposal put forth in Lowenstamm (1999) achieves exactly this goal. Subject–object sequences, indefinite genitive constructions and nominal compounds are represented as in (13a–c) respectively:[28]

(13) Phonological representation of word-initial gemination and word-medial gemination
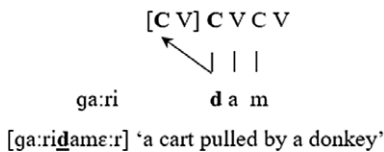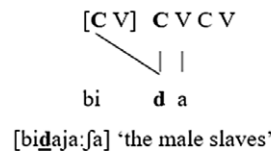
a. Subject-object: *saxafi danjiri(hii)*

    [C V] C V C V

    saħafi   **d** a n

    [saħafi**d**antʃiri:] 'journalist ambassador'

b. Indefinite genitives: *jeeni dameer*

    [C V] C V C V

    d͡ʒeːni   **d** a m

    [d͡ʒeːni**d**amɛːr] 'front leg of a donkey'

c. Nominal compounds: *gaari-dameer*

    [C V] C V C V

    ɡaːri   **d** a m

    [ɡaːri**d**amɛːr] 'a cart pulled by a donkey'

d. Word-medial geminates: *biddayaasha*

    [C V] C V C V

    bi   **d** a

    [bi**d**ajaːʃa] 'the male slaves'

Word-initial gemination is accounted for as follows: in all cases, Noun2 is preceded by an additional timing unit (an empty skeletal site), on which the initial stop propagates. The representation of the resulting stop is identical to that of a lexical geminate, represented in (13d).[29] This analysis clearly encodes the fact that word-initial singletons and geminates pattern alike at the temporal level: temporal properties are represented at the skeletal level; like geminates, word-initial singletons are longer than word-medial singletons. (By contrast, non-temporal parameters, e.g. release amplitude could be considered to encode domain-initial strengthening. Further research on the fine articulatory properties of voiced stops in Somali, e.g. the amount of linguo-palatal contact, is needed to draw firmer conclusions on this issue.)

We conclude this discussion by a brief outlook on the predictions made by our proposal. Initial [CV] is assumed to be present in front of major categories, only. This makes predictions on the realization of the initial voiced stop of minor categories on the one hand, and of major categories other than the noun on the other hand. Consider first the minor categories, exemplified by the determiner. The underlying form of the Somali determiner is *-ta* (feminine)/*-ka* (masculine). Voiceless stops are voiced in intervocalic position, resulting in *-da/-ga* (Armstrong 1934; Bell 1953: 12; Saeed 1999: 28ff.; Barillot 2002: 232ff.). Since

---

[28] CV stands for a sequence of a consonantal slot and a vocalic slot. Note that our proposal does not hinge on the specific CV-framework: it can straightforwardly be adapted into any autosegmental phonology framework distinguishing between the skeletal level and the segmental level, very much in the spirit of Guerssel's analysis of Berber verbal morphology.

[29] Initial CV is identified if the phonotactic constraints are met, i.e. in intervocalic position, only. Otherwise, gemination would lead to a banned CCC cluster.
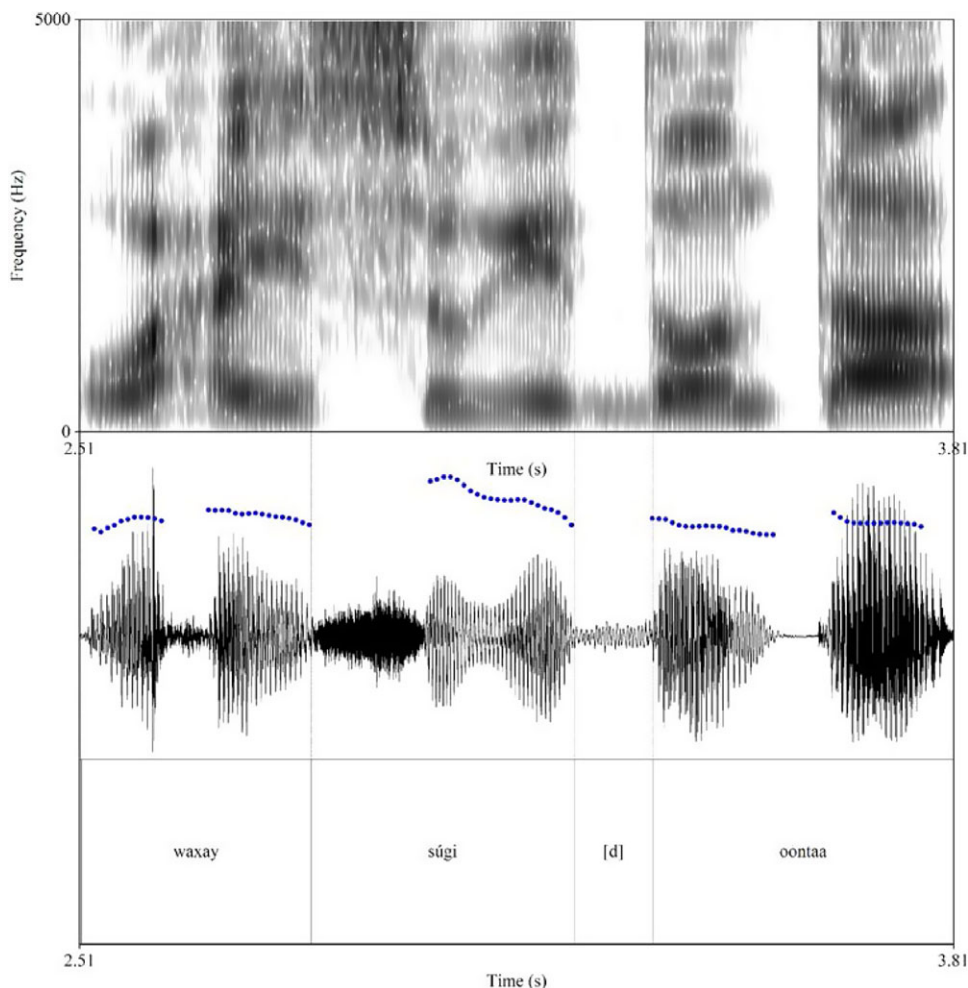
Figure 11 (Colour online) Waveform, spectrogram and pitch curve (speckles above the waveform) of the infinitive-verb sequence *waxay sugi doontaa* 'she will wait'. /d/ clearly exhibits a release burst; it is realized as the fully voiced stop [d] with a closure duration of 117 ms.

the determiner is a minor category, we expect the initial d/g of the determiner to surface as an approximant, not as a stop. This prediction is borne out: all instances of Noun-Determiner sequences in our recordings confirm the findings of Armstrong (1934) and Farnetani (1981), e.g. /maga:lo+ta/ 'city+the' → /maga:láda/ → [maɣa:láð̞a] 'the city'.[30] Finally, we predict

---

[30] Note that our approach accounts for the Somali facts without postulating a recursive structure for the prosodic word. By contrast, consider for instance the approach recently defended by Green & Morrison (2016). They assume that each accented suffixed determiner in Somali constitutes a ω, which is grouped into a higher ω in a recursive structure, exactly like Noun2 in Noun1 Noun2 compounds: see their (26), [guriɣaːɣěː] 'house–your–which' analysed as ((gúri)ω (gáa)ω (geé)ω)ω. Noun–Determiner sequences in Somali are thus expected to behave like nominal compounds. However, the initial voiced stop of the determiner is realized as an approximant, while the initial voiced stop of Noun2 in a Noun1 Noun2 compound is realized as a stop. In order to 'save' this line of reasoning, one could suggest that each noun in a compound constitutes a domain which is lower than ω. However, this option faces two major problems: first, it cannot account for the fact that the initial stop of N2 in a N1 N2 compound behaves

word-initial gemination to apply with all major categories, not only noun-initially. We thus expect it to apply verb-initially, too. Again, this prediction is borne out: a first survey of the noun–verb compounds and infinitive-verb sequences recorded in our corpus shows that verb-initial /b d g/ surfaces as a stop, not as an approximant. This is exemplified in Figure 11 with the sequence *súgi doontaa* 'wait will' → '(she) will wait', which is realized as [. . .idoːn. . .], *[. . .iðoːn. . .].

5 Conclusion

This article explored the acoustic properties of Somali intervocalic singleton and geminate voiced stops through a production experiment. More specifically, we sought to determine the role of the relevant temporal and non-temporal acoustic correlates in the realization of Somali geminate and singleton /b d g/. The first issue was that of the contrast between word-internal /b d g/ and /bb dd gg/; the second issue concerned the realization of word-initial /b d g/, which exhibits various peculiarities that are reminiscent of gemination or of domain-initial strengthening.

Our results show that word-internal singletons are consistently realized as open approximants (with no release burst, but with a formant structure and a high level of energy). They contrast with geminates, which are consistently realized as fully voiced stops, with a strikingly short closure duration (and a low closure amplitude). We conclude that the opposition between singleton and geminate voiced stops is primarily realized as the MANNER CONTRAST approximant [β ð̞ ɣ] vs. short stop [b d g]. We propose an analysis that reconciles the acoustic properties of intervocalic geminates (short duration) with their phonological behaviour (two skeletal slots).

Concerning the word-initial context, our results establish that word-initial voiced singleton stops and word-medial geminates share the same closure duration, release burst duration, and vowel duration within the Prosodic Word. They also have a similar closure amplitude, and voicing properties. Moreover, the acoustic properties of word-initial singleton stops remain constant, and do not depend on their position in the prosodic hierarchy. These results lead us to propose that there are only two categories of voiced consonants in Somali: word-medial singletons, on the one hand (approximants), and word-medial geminates and word-initial singletons, on the other hand (short voiced stops). Based in particular on the temporal similarities between word-initial singleton voiced stops and medial geminates, we propose that word-initial stops have the same phonological representation as word-medial geminates, with two skeletal slots. Word-initial strengthening in Somali is essentially word-initial LENGTHENING.

---

like a lexical geminate; second, the only motivation for this lower domain would be the observation that the initial consonant of N2 is strengthened, i.e. the proposal becomes circular.

Appendix

Mean values for temporal (Table A1) and non-temporal (Table A2) parameters.

Table A1 Mean values for closure duration, release duration and vowel duration for each consonant and each context. Durations are given in milliseconds with standard deviations in parentheses. Contexts: Cmp = Compounds, Gen = Genitive constructions, Ind = subject–object sequences, LexC/CC = Lexical singletons/geminates.

	Context				
Duration	Cmp	Gen	Ind	LexC	LexCC
Closure duration					
b	77.5 (9.9)	81.2 (12.3)	76.1 (10.7)	43.1 (9.9)	88.7 (27.6)
d	84.5 (12.0)	80.4 (15.6)	83.9 (11.1)	45.1 (12.7)	77.2 (21.4)
g	75.4 (13.7)	70.3 (16.9)	72.0 (16.4)	43.1 (8.1)	72.7 (13.6)
Release duration					
b	9.6 (2.8)	8.8 (2.8)	9.9 (4.3)	DNA	8.8 (3.3)
d	11.7 (3.5)	12.0 (4.2)	11.6 (2.6)	DNA	11.7 (4.0)
g	20.3 (7.2)	21.9 (7.0)	17.9 (6.2)	DNA	17.8 (5.0)
Vowel duration					
b	53.0 (16.2)	66.6 (37.0)	79.5 (16.8)	51.5 (12.5)	50.5 (14.7)
d	59.5 (14.7)	71.5 (15.2)	76.9 (15.9)	62.2 (13.7)	55.9 (13.9)
g	52.6 (11.9)	70.5 (21.6)	81.9 (13.9)	57.4 (11.5)	63.4 (14.1)

Table A2 Mean values for closure amplitude and release amplitude for each consonant and each context. Values for standard deviation appear in parentheses. Contexts: Cmp = Compounds, Gen = Genitive constructions, Ind = subject–object sequences, LexC/CC = lexical singletons/geminates.

	Context				
Amplitude	Cmp	Gen	Ind	LexC	LexCC
Closure amplitude					
b	21.6 (7.7)	18.8 (14.8)	16.6 (7.1)	47.4 (21.1)	24.0 (15.6)
d	22.3 (9.2)	21.2 (8.9)	26.4 (9.9)	41.5 (19.9)	33.4 (19.4)
g	23.3 (9.9)	17.9 (8.1)	24.7 (9.9)	42.3 (22.6)	23.5 (11.4)
Release amplitude					
b	24.24 (15.95)	26.60 (18.00)	23.14 (10.67)	DNA	18.55 (11.70)
d	17.54 (8.76)	22.40 (11.24)	25.23 (12.01)	DNA	21.58 (12.30)
g	14.73 (9.24)	18.24 (11.55)	20.37 (12.02)	DNA	13.91 (7.06)

References

Abdirachid, Mohamed Ismail. 2011. *Dialectologie du somali : problématiques et perspectives*. Thèse de doctorat, INALCO Paris.

Abdullahi, Mohamed Diriye. 2001. *Le somali, dialectes et histoire*. Thèse de doctorat, Université de Montréal.

Andrzejewski, B. W. 1955. The problem of vowel representation in the Isaaq dialect of Somali. *Bulletin of the School of Oriental and African Studies* 17(3), 568–580.

Andrzejewski, B. W. 1964. *The declensions of Somali nouns.* London: University of London, School of Oriental and African Studies.

Andrzejewski, B. W. 1979. *The case system of Somali*. London: University of London, School of Oriental and African Studies.

Armstrong, Lilias E. 1934. The phonetic structure of Somali. *Mitteilungen des Seminars für orientalische Sprachen zu Berlin* 37(3), 116–161.

Baayen, Rolf H., Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4), 390–412.

Banti, Giorgio. 1988. Two Cushitic systems: Somali and Oromo nouns. In Harry van der Hulst & Norval Smith (eds.), *Autosegmental studies on pitch accent*, 11–50. Dordrecht: Foris.

Barillot, Xavier. 2002. *Morphophonologie gabaritique et information consonantique latente en somali et dans les langues est-couchitiques.* Thèse de doctorat, Université Paris 7 – Denis Diderot.

Barillot, Xavier & Philippe Ségéral. 2005. On phonological processes in the "3rd" conjugation in Somali. *Folia orientalia* 41, 115–131.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67(1), 1–48.

Bell, Christopher R. V. 1953. *The Somali language*. London: Longmans, Green & Co.

Boersma, Paul & David Weenink. 2019. Praat: Doing phonetics by computer. http://www.fon.hum.uva.nl/praat/ (accessed 21 December 2019).

Cardona, Giorgio Raimondo. 1981. Profilo fonologico del somalo. In Cardona & Agostini (eds.), 3–26.

Cardona, Giorgio Raimondo & Francesco Agostini (eds.). 1981. *Fonologia e Lessico* (Studi Somali 1). Roma: Ministero degli Affari Esteri, Dipartimento per la Cooperazione allo Sviluppo.

Cauquil, Laurent & Sylvie Combes. 2019. Modèles mixtes – guidelines. http://www.nathalievialaneix.eu/doc/html/guidelines_modele_mixte.html#comment_ (accessed 4 February 2021).

Cho, Taehong. 2011. Laboratory Phonology. In Nancy C. Kula, Bert Botma & Kuniya Nasukawa (eds.), *The Continuum companion to phonology*, 343–368. London & New York: Continuum.

Cho, Taehong & Patricia A. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29(2), 155–190.

Cho, Taehong & Patricia A. Keating. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37(4), 466–485.

Cohn, Abigail C., William H. Ham & Robert J. Podesva. 1999. The phonetic realization of singleton–geminate contrasts in three languages of Indonesia. In John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashlee C. Bailey (eds.), *14th International Congress of Phonetic Sciences* (ICPhS XIV), San Francisco, CA, 587–590.

Cunnings, Ian. 2012. An overview of mixed-effects statistical models for second language researchers. *Second Language Research* 28(3), 369–382.

Downing, Laura J. & Morgan Nilsson. 2019. Prosodic restructuring in Somali nominal constructions. In Samson Lotven, Silvina Bongiovanni, Phillip Weirich, Robert D. Botne & Samuel Gyasi Obeng (eds.), *African linguistics across the disciplines: Selected papers from the 48th Annual Conference on African Linguistics*, 125–142. Berlin: Language Science Press.

Edmondson, Jerold A., John H. Esling & Jimmy G. Harris. 2003. Supraglottal cavity shape, linguistic register, and other phonetic features of Somali. Ms., https://api.semanticscholar.org/CorpusID:34038524.

Ehret, Christopher & Mohamed Nuuh Ali. 1984. Soomaali classification. In Thomas Labahn (ed.), *Proceedings of the Second International Congress of Somali Studies*, vol. 1: *Linguistics and Literature*, 201–270. Hamburg: Helmut Buske.

Farnetani, Edda. 1981. Dai tratti ai parametri: introduzione all'analisi strumentale della lingua somala. In Cardona & Agostini (eds.), 27–108.

Figueroa Candia, Mauricio A. 2016. *Lenition in the production and perception of Chilean Spanish approximant consonants: Implications for lexical access models*. Ph.D. dissertation, University College London.

Flege, James E. 1982. Laryngeal timing and phonation onset in utterance-initial English stops. *Journal of Phonetics* 10, 177–192.

Fougeron, Cécile & Patricia A. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America* 101(6), 3728–3740.

Green, Christopher R. & Michelle E. Morrison. 2016. Somali wordhood and its relationship to prosodic structure. *Morphology* 26, 3–32.

Grosjean, François & Alain Deschamps. 1972. Analyse des variables temporelles du français spontané. *Phonetica* 26, 129–156.

Guerssel, Mohand. 1992. The phonology of Berber derivational morphology by affixation. *Linguistic Analysis* 22, 3–60.

Gussenhoven, Carlos. 2004. *The phonology of tone and intonation* (Research Surveys in Linguistics). Cambridge: Cambridge University Press.

Hamzah, Mohd Hilmi, Janet Fletcher & John Hajek. 2012. An acoustic analysis of release burst amplitude in the Kelantan Malay singleton/geminate stop contrast. In Felicity Cox, Kathrin Demuth, Susan Lin, Kelly Miles, Sallyanne Palethorpe, Jason Shaw & Ivan Yuen (eds.), *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, 85–88. Canberra: ASSTA.

Hamzah, Mohd Hilmi, Janet Fletcher & John Hajek. 2016. Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay. *Journal of Phonetics* 58, 135–151.

Hassan, Mohamed Moallin. 1994. *Aspects de la phonologie et de la morphologie du somali*. Thèse de doctorat, Université de Nice-Sophia Antipolis.

Hirst, Daniel & Albert Di Cristo (eds.). 1998. *Intonation systems: A survey of twenty languages*. Cambridge: Cambridge University Press.

Hock, Hans H. 1991. *Principles of historical linguistics*, 2nd edn. Berlin: Mouton de Gruyter.

Hyman, Larry M. 1981. Tonal accent in Somali. *Studies in African Linguistics* 12(2), 169–201.

Jaeger, Jeri J. 1978. Speech aerodynamics and phonological universals. In Jeri J. Jaeger, Anthony C. Woodbury, Farrell Ackerman, Christine Chiarello, Orin D. Gensler, John Kingston, Eve E. Sweetser, Henry Thompson & Kenneth W. Whistler (eds.), *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society* (BLS 4), 312–329. Berkeley, CA: Berkeley Linguistics Society.

Keating, Patricia A. 2006. Phonetic encoding of prosodic structure. In Jonathan Harrington & Marija Tabain (eds.), *Speech production*, 167–83. New York: Psychology Press.

Keating, Patricia [A.], Taehong Cho, Cécile Fougeron & Chai-Shune Hsu. 2003. Domain-initial strengthening in four languages. In John Local, Richard Ogden & Rosalind Temple (eds.), *Laboratory Phonology VI: Phonetic interpretation*, 145–163. Cambridge: Cambridge University Press.

Kingston, John, Shigeto Kawahara, Della Chambless, Daniel Mash & Eve Brenner-Alsop. 2009. Contextual effects on the perception of duration. *Journal of Phonetics* 37(3), 297–320.

Krakow, Rena A., Fredericka Bell-Berti & Q. Emily Wang. 1994. Supralaryngeal declination: Evidence from the velum. In Fredericka Bell-Berti & Lawrence J. Raphael (eds.), *Producing speech: A Festschrift for Katherine Safford Harris*, 333–353. Woodbury, NY: AIP Press.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. *lmerTest* package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(1), 1–26.

Ladd, D. Robert. 2008. *Intonational phonology*, 2nd edn. Cambridge: Cambridge University Press.

Ladefoged, Peter & Ian Maddieson. 1995. *The sounds of the world's languages*. Oxford Cambridge, MA: Blackwell.

Lahiri, Aditi & Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16, 327–338.

Lamberti, Marcello. 1986. *Die Somali-Dialekte: eine vergleichende Untersuchung*. Hamburg: Helmut Buske.

Le Gac, David. 2001. *Structure prosodique de la focalisation : le cas du somali et du français*. Thèse de doctorat, Université Paris 7 – Paris Diderot.

Le Gac, David. 2003a. Marques prosodiques de la focalisation contrastive en somali. In Anne Lacheret-Dujour (ed.), *Fonction et moyens d'expression de la focalisation à travers les langues* (Mémoires de La Société Linguistique de Paris – Nouvelle Série XIII), 49–80. Leuven: Peeters.

Le Gac, David. 2003b. Tonal alternations in Somali. In Jacqueline Lecarme (ed.), *Research in Afroasiatic grammar II* (Current Issues in Linguistic Theory 241), 287–304. Amsterdam & Philadelphia, PA: John Benjamins.

Le Gac, David. 2018. Two patterns of tone lowering in Somali. In Katarzyna Klessa, Jolanta Bachan, Agnieszka Wagner, Maciej Karpiński & Daniel Śledziński (eds.), *Proceedings of the 9th International Conference on Speech Prosody 2018*, 99–103. Poznań, Poland: ISCA.

Lecarme, Jacqueline. 1999. Focus in Somali. In Georges Rebuschi & Laurice Tuller (eds.), *The grammar of focus*, 275–310. Amsterdam & Philadelphia, PA: John Benjamins.

Local, John & Adrian P. Simpson. 1999. Phonetic implementation of geminates in Malayalam nouns. *Proceedings of the 14th International Congress of Phonetic Sciences* (ICPhS XIV), San Francisco, CA, 595–598.

Lowenstamm, Jean. 1996. CV as the only syllable type. In Bernard Laks (ed.), Current trends in phonology, 419–441. Manchester: ESRI.

Lowenstamm, Jean. 1999. The beginning of the word. In John R. Rennison & Klaus Kühnhammer (eds.), *Phonologica 1996*, 153–166. The Hague: Holland Academic Graphics.

Martínez-Celdrán, Eugenio. 2004. Problems in the classification of approximants. *Journal of the International Phonetic Association* 34(2), 201–210.

Martínez-Celdrán, Eugenio & Xosé L. Regueira. 2008. Spirant approximants in Galician. *Journal of the International Phonetic Association* 38(1), 51–68.

Nilsson, Morgan. 2017. Somali as a pluricentric language. Presented at the 5th World Conference on Pluricentric Languages and their non-dominant Varieties, 13–15 July, University of Mainz, Germany. http://morgannilsson.se/ppt/2017-07-13-Mainz.pdf.

Nilsson, Morgan. 2018. Somali as a pluricentric language: Corpus-based evidence from schoolbooks. In Rudolf Mur & Benjamin Meisnitzer (eds.), *Pluricentric languages and non-dominant varieties worldwide: New pluricentric languages – old problems*, 75–91. Frankfurt a. M.: Peter Lang.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Ohala, John J. & Carol J. Riordan. 1979. Passive vocal tract enlargement during voiced stops. *The Journal of the Acoustical Society of America* 65(S1), S23–S23.

Orwin, Martin. 1994. *Aspects of Somali phonology*. Ph.D. dissertation, University of London, School of Oriental and African Studies.

Orwin, Martin. 1995. *Colloquial Somali*. London & New York: Routledge.

Puglielli, Annarita (ed.). 1984. *Aspetti morfologici, lessicali e della focalizzazione* (Studi Somali 5). Roma: Ministero degli Affari Esteri, Dipartimento per la Cooperazione allo Sviluppo.

R Core Team. 2019. R: The R project for statistical computing. https://www.r-project.org/ (accessed 12 February 2020).

Ridouane, Rachid. 2007. Gemination in Tashlhiyt Berber: An acoustic and articulatory study. *Journal of the International Phonetic Association* 37(2), 119–142.

Ridouane, Rachid. 2010. Gemination at the junction of phonetics and phonology. In Cécile Fougeron (ed.), *Laboratory Phonology 10: Phonology and phonetics*, 61–91. Berlin & New York: Mouton de Gruyter.

Saeed, John I. 1982. Central Somali: A grammatical outline. *Afroasiatic Linguistics* 8(2), 1–43.

Saeed, John I. 1999. Somali (The London Oriental and African Library). Amsterdam & Philadelphia, PA: John Benjamins.

Ségéral, Philippe & Tobias Scheer. 2001. La Coda-Miroir. *Bulletin de la Société Linguistique de Paris* 96(1), 107–152.

Solé, Maria-Josep. 2018. Articulatory adjustments in initial voiced stops in Spanish, French and English. *Journal of Phonetics* 66, 217–241.

Titze, Ingo R. 1988. The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America* 83(4), 1536–1552.

Turk, Alice E. & Laurence White. 1999. Structural influences on accentual lengthening in English. *Journal of Phonetics* 27(2), 171–206.

Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America* 91(3), 1707–1717.

Zorc, R. David & Madina M. Osman. 1993. *Somali–English dictionary with English index*, 3rd edn. Kensington, MD: Dunwoody Press.