Migration Letters

Volume: 20, No: 8, pp. 1225-1238

ISSN: 1741-8984 (Print) ISSN: 1741-8992 (Online)

www.migrationletters.com

The Voice-Onset Time of Arabic Loanwords in Swahili

Mohammad Alsamaani[1]

Abstract

In their recent survey of loanword adaptation, Paradis and LaCharité (2011) note that voice-onset time (VOT) is useful in distinguishing between phonetic and phonological approximation. The present study considers the role of VOT in Arabic-adapted Swahili words, focusing on initial stops. The results are mixed. On the one hand, it is found that voiceless stops have lower VOT values in Swahili than in Arabic, supporting the idea that Swahili lacks aspiration for Arabic loans, further suggesting that the borrowings have a phonological basis. On the other hand, the voiced stops in both languages show very similar negative values, which substantiate the influence of phonetics on adaptations. The VOT facts reported in this study are valuable independently of adaptation phenomena, as relatively little is known about VOT in Arabic and especially Swahili, in spite of a long tradition of VOT studies (see especially Lisker & Abramson 1964, and Cho & Ladefoged 1999).

Keywords: *Arabic, Swahili, Male, VOT, Computational Linguistics.*

1. Introduction

Cross-cultural interaction has brought Swahili (Kiswahili), from the Bantu branch of the Niger-Congo family of languages, into contact with Arabic, from the Semitic branch of Afro-Asiatic (Trask 2007: 235-236). Swahili is especially dominant in Tanzania and Kenya, where it is used for everyday interactions. There is a scant literature concerning the sound inventory of Swahili, but descriptions are limited and unclear about whether they are broadly phonological or narrowly phonetic. For instance, Velupillai (2012: 66) noted that in the "belt across Africa, south of the Sahara," Swahili is one of the 11.8% of languages catalogued in UPSID[2] that feature voiced implosives. Indeed, some linguists who have discussed Swahili assert that it makes a phonological distinction between plosives and implosives (Tucker & Ashtone 1942; Polomé 1967; Hayward et al. 1989; Kharusi 1995). Others who have briefly studied Swahili stops have paid little attention to their voicing, focusing instead on the distinction between aspiration and non-aspiration (Engstrand & Lodhi 1985; Frankl 1991; Lodhi 2003; Alsamaani 2021).

According to Polomé (1967: 38-40), there are four laryngeal categories of Swahili stops across people with average education: voiced plosives and implosives, voiceless aspirated, and unaspirated plosives. His observations presented a combination of twelve plosive and implosive sounds: /p/, /pʰ/, /b/, /ɓ/, /t/, /tʰ/, /d/, /ɗ/, /k/, /kʰ/, /g/ and /ɠ/. He

[1] Department of English Language & Translation, College of Arabic Language & Social Studies, Qassim University, Buraydah, Saudi Arabia. E-mail: msamaany@qu.edu.sa

[2] The UCLA Phonological Segment Inventory Database (Maddieson 1981).

noted that the aspiration contrast is in decline in the South of the Swahili-speaking area (e.g., Zanzibar and Tanzania). This is expected for Polomé, as the contrast in aspiration originated in the North (e.g., Kenya). For instance, the initial stop of paka meaning 'cat' is unaspirated in Zanzibar, whereas it is aspirated in the North, e.g. Mombasa p^haka 'cat'. The word paka meaning 'smear' is unaspirated in both regions. That is, Southern varieties of Swahili have merged aspirated [p^h] with [p].

Engstrand and Lodhi (1985) worked with 20 hours of recordings of 20 native and 15 non-native[3] Swahili speakers from Zanzibar, Tanga, Darussalam and Mombasa. After having listened to the audio files of all participants, Engstrand and Lodhi drew conclusions that undermined their confidence in Polomé's work and caused them to question the credibility of his informants.

Of particular interest is Engstrand and Lodhi's observation (see also Frankl 1991) that Arabic borrowed words in Swahili are exempt from the influence of aspiration, with only a few exceptions. This fact raises several questions: Is aspiration in Swahili the result of borrowings from other bordering languages and local dialects that it dealt with differently from Arabic? If this assumption is proven true, then what about implosives? Are they exceptions to words of purely Bantu origin? The aforementioned discrepancy needs both phonetics and phonology to examine the loanwords from Arabic. As in Frankl (1991: 369), on the one hand, the velar voiceless stop /k/ from Arabic is mostly adopted as unaspirated, much like what can be found in the word kahawa, or "coffee" (Ar. qahwah). On the other hand, the word k^hurasa, "notebook" (Ar. kurrāsah), is adapted with aspiration. Frankl's explanation for k^hurasa, is manifested in its deep immersion into the Swahili language. Has the source sound influenced this asymmetry in the receptor language adaptation /q/ to /k/ vs. /k/ to /k^h/?

Accordingly, linguists need a practical tool for judging the common rule for initial stops of Swahili borrowings from Arabic. Is it phonetically or phonologically motivated?

In their defence of the phonological stance, Paradis and LaCharité (2011: 755-756) point to phonetic differences in the voice-onset time (VOT) of bilabial voiced and voiceless stops in both English and French. They note that in English, /p/ has a positive long lag VOT, whereas /b/ has a positive short lag one. Strikingly, /p/ in French falls within the same VOT region as /b/ in English, whereas /b/ in French has a negative lead VOT (see Ladefoged and Johnson 2011: 154; Netelenbos et al. 2015). Thus if English /b/ were to be adapted into French phonetically, it ought to become /p/. By contrast, the phonological adaptation would maintain the same category, causing /b/ to retain its phonological features during borrowing. That is, French /b/ ought to become English /b/ on the phonological account, and this is indeed what happens.

On the other hand, Ito and Kenstowicz (2013) illustrate the importance of phonetics in their study of Japanese loanwords in Korean. In particular, they demonstrated a diachronic shift in the adaptation of stops due to ongoing changes in phonetics. According to Ito and Kenstowicz's analysis, in the Classical period the Korean aspirated/tense stops were categorically different from the voiceless stops in Japanese, so the Japanese ones were adapted as lax. Over time, however, the Korean aspirated stops and Japanese voiceless stops have moved closer, resulting in a near merger along the VOT continuum. Therefore, the contemporary Korean adaptation of the Japanese voiceless stops has biased towards aspirated over lax.

Stops are not the only features that may show alternation in the VOT; a study by Port and Rotunno (1979) claims that the vowel following the consonant has some impact on the

[3] Native speakers are defined as those who have no Arabic influence and non-natives are defined as second language speakers.

VOT duration[4]. Their experiment targeted six combinations of lax/tense vowels (i/ī, u/ū, ʌ/ɑ) after voiceless stops /p/, /t/, and /k/. It was concluded in this article that the velar has the highest VOT followed by alveolar and then bilabial. Moreover, tense vowels are believed to have affected the duration of the preceding stops positively more than the lax ones. Klatt (1975: 686) argued against this claim in an earlier study by reporting that “the VOT was longer before sonorants and high vowels than before mid and low vowels.” Although the effect of vowels is not the main interest of this study, it is important to keep in mind the variables that have some impact on VOT representation.

Revisiting the Swahili k^hurasa and kahawa above, the VOT of these words and many others need to be phonetically measured in order to determine which one is the successor: phonetics or phonology. I hypothesize that the uvular /q/ in Arabic has a short lag voice, which led to its adaptation as a non-aspirated velar /k/ in Swahili. Conversely, the velar /k/ in Arabic may have presented some sort of breathy voice, resulting in an aspirated /k^h/ in Swahili. In her analysis of English loanwords in Korean, Kang (2010) stressed the consideration of all factors that would introduce the phonological adaptation, namely the influence of phonetics, and in her case, the process of borrowing loanwords. Adopting Kang’s neutral view of loanword adaptation, if we buy the idea that Arabic stops include no aspiration, then the diminishing of [the spread glottis] feature in the dialect of Zanzibarites[5] as claimed by Polomé (1967) is a result of the phonological effect of Arabic. As a part of the discussion, we will attempt to clarify whether this assumption holds true or not.

The aim of this study is to phonetically examine some selected Swahili words of Arabic origin. The words’ VOT of the initial stops will be measured and compared to the source equivalents. The list under investigation illustrates that category mapping of the stops match the donor consonantal inventory except for the Arabic uvular /q/ and alveolar emphatic /$t^ʕ$/ as they were borrowed as /k/ and /t/, respectively. The transition from /q/ and /$t^ʕ$/ to /k/ and /t/ will be discussed later. Section 2 posits a theory for this phenomenon by providing a background and context for the VOT of stops universally, then comments on Arabic and Swahili. In section 3, we discuss the database. Section 4 is the result and ensuing discussion. The last section, 5, is the conclusion.

2. Background

The background is meant to bring about some early studies dealt with voice onset-time of stops in some languages of the world. It also discusses the available literature in concern with VOT in Arabic and Swahili.

2.1 VOT Cross-Linguistically

VOT is a phonetic tool used to observe the differences between consonants among languages of the world. Though this concept was introduced earlier, it was not until the work of Lisker and Abramson (1964) when VOT gained traction in the field. In the summary at the end of their article, they referred to their procedure of measuring VOT as “the time interval between the burst that marks release and the onset of periodicity that

[4] In their recent work on English spontaneous speech in word initial stops, Nakai & Scobbie (2016: 20) report that the frequency of a stop before any given vowel has an influence over the VOT length (an increase for voiced stops but a decrease for the voiceless ones.) Further, “The more frequent content and function words were, the more likely they were to have a VOT that fell on the opposite side of the optimal category boundary, producing an overlap between voiced vs. voiceless VOT distributions” (p.22)

[5] The residents of the Island in general are called Zanzibaries, but those who live in Zanzibar city are referred to as Zanzibarites (Engstrand & Lodhi 1985: 180). It is worth mentioning here that Dr. Lodhi was born and raised in Zanzibar before immigrating to Sweden.

reflects laryngeal vibration" (p. 220). They studied 11 languages, which they divided into three groups depending on their stop categories: "voicing, aspiration and force of articulation." Since the contrast between the languages was mostly dependent on the phoneme categories of each language, English was among the languages that exhibited a two-way contrast in voicing. Due to being word-initial in the experiment, voiceless stops were assigned a long lag VOT (i.e. aspiration). For more information, see table 1 below. The mean values for 6 of the 11 languages where English belongs are -100, +10 and +75 ms.[6] Another point that raises some concern in this regard is the small number of subjects used in this study; each one of the 7 out of the 11 languages had only 1 informant (i. e. Dutch, Tamil, Cantonese, Eastern Armenian, Korean, Hindi and Marathi).

Table 1. A summary for VOT mean of initial stops in (ms) from Lisker & Abramson (1964).[7]

Lang. / Stop	Dutch	Puerto Rican Spanish	Hungarian	Tamil	Cantonese	English[8]	Eastern Armenian	Thai	Korean[9]	Hindi	Marathi
/p/	10	4	2	12	9		3	6	7/18	13	11
/t/	15	9	16	8	14		15	9	11/25	15	10
/k/	25	29	29	24	34		30	25	19/47	18	24
/pʰ/					77	58	78	64	91	70	76
/tʰ/					75	70	59	65	94	67	65
/kʰ/					87	80	98	100	126	92	87
/b/	-85	-138	-90	-74		1	-96	-97		-85	-117
/d/	-80	-110	-87	-78		5	-102	-78		-87	-111
/g/		-108	-58	-62		21	-115			-63	-116

In a later study on 18 endangered languages by Cho and Ladefoged (1999), they utilized recordings from the UCLA phonetic lab. Four positive categories are distinguished for the VOT of aspirated and non-aspirated stops (unaspirated, slightly aspirated, aspirated, and highly aspirated). Therefore, the article has shown that the range for unaspirated VOT values begin around 30 ms. 50 ms is the starting zone of a lesser aspiration, and from around 90 ms is the range for aspiration. Furthermore, the point on the continuum at which aspiration reaches its climax occurs in languages with around and over 120 ms. The two authors have looked at places of articulation of "voiceless" stops: bilabial, dental, alveolar, velar, and uvular. The highest records for VOT values are as follows[10]: (a) labial is by Khonoma Angami at 83 ms, (b) dental is by Tlingit at 120 ms, (c) alveolar is by Navajo at 130 ms, (d) velar is by Navajo at 154 ms, (e) uvular is by Tlingit at 128 ms. The lowest unaspirated VOT values are recorded in the following: (a) labial is by Khonoma Angami at 10 ms, (b) dental is by Khonoma Angami at 9 ms, (c) alveolar is by

[6] It should be noted that it is not the norm for English to have a voicing lead in its voiced stops, yet it was categorized as such because one of the participants was responsible for most of the negative voicing.

[7] Stops in the table are voiced, and voiceless aspirated and non-aspirated.

[8] The mean of /b/, /d/, /g/ of English excludes the coded (TR) speaker because of the negative lead responses.

[9] They distinguished between the two values by assigning an apostrophe symbol /'/ to (18, 25 and 47). This would be an intermediate stage between plain and aspirated stops.

[10] As some languages show aspiration/non-aspiration contrast, they may appear twice in the comparison.

Navajo at 6 ms, (d) velar is by Khonoma Angami at 20 ms, (e) uvular is by Hupa at 27 ms.[11]

Full agreement on the VOT categorization of each specific language has not been reached. Commenting on the overlapping results of Lisker and Abramson's study vs. Cho and Ladefoged's as mentioned above, Chao and Chen (2008) stated that English and Chinese are both aspirated as in Lisker and Abramson's data. On the contrary, they further reported that the degree of aspiration in Cho and Ladefoged's work would assign English and Chinese different levels of aspiration. In this case, Chinese with VOT means of 77, 75, and 87 ms[12] would be closer to the aspirated VOT continuum than English (58, 70 and 80 ms).

In order to test the assumption of Lisker and Abramson, who assigned English a two-way contrast, Chao and Chen (2008) conducted their own study by comparing the speech of 11 Mandarin speakers with 4 speakers of British English.[13] Their conclusion implied that both English and Mandarin occupy the same range for non-aspirated stops along the VOT spectrum; however, aspirated languages depart from the result of the '60s study by assuring that the level of aspiration of the two languages is asymmetrical in nature and both languages should be ascribed a different degree of aspiration. Moreover, by building on the ideas of Cho and Ladefoged (1999) above, the authors finalized the outcomes of their research by backing the four categories proposed in the 1999 study. Thus, because of the higher mean values of Mandarin (82, 81, and 92 ms) versus those of English (62, 73, and 86 ms), English should be classified as aspirated, and Mandarin, highly aspirated.

As a supporting argument for the effect of backward directionality on the lengthening of VOT values in English, Nakai & Scobbie (2016: 13) back the result of Rotunno (1979). They show that the shortest median value is assigned for voiced labial stops by 2 ms, yet alveolars and velars have 6 ms and 17 ms respectively. Nakai & Scobbie also find that the boundary location in milliseconds for voiceless stops is 16 for labilas, 24 for alveolars and 27 for velars.

As a closing point of this section, it should be mentioned that the historical phonetics/phonology of adaptation has interesting remarks on the process of introducing VOT to the borrowing community. In the case of Ito and Kenstowicz's (2013) analysis of Korean adaptation from Japanese and the way such reception of the stops of newly introduced words may change over time, the presumed deterioration of aspiration in the South and Zanzibar in particular might be best accounted for by looking into the historical part of adaptation. It should be noted that while this is not under the scope of this study, it is a strong implication for future work. Hence, closed and/or isolated communities may have preserved the stops of classical loans faithfully.

2.2 VOT in Arabic.

It is worth noting that we should know the sound inventory of stops in Standard Arabic (SA) before delving into the literature that dealt with the VOT of plosives in Arabic (see table 2). First, Arabic misses the bilabial voiceless stop /p/ in its plosive's system. It only has the voiced bilabial stop /b/. The place of articulation of the alveolar has four consonantal sounds, two plain /t/, /d/ and two emphatics, /tˤ/, /dˤ/. Besides, Arabic has the velar /k/ with no counterpart, the uvular /q/ and the glottal /ʔ/.

[11] Khonoma Angami: Nagaland, India, Tlingit: Alaska, U.S.A., Navajo: New Mexico, Hupa: CA, U.S.A (Cho & Ladefoged, 1999: 216-217).
[12] The sequence of numbers between parentheses refers first to bilabial, and then alveolar followed by velar.
[13] I adapt their result of the English participants without the fourth because she shows some sign of bilingualism.

Table 2. SA Consonants after (Batais 2013; Bin-Muqbil 2006).

Stops	Emphatic stops	Fricatives	Emphatic Fricatives	Affricates	Nasal	Trill	Glides	Lateral	Emphatic Lateral
b, d, t, k, q, ʔ	tˤ, dˤ	f, θ, ð, s, z, ʃ, x, ɣ, ħ, ʕ, h	ðˤ, sˤ	dʒ	m, n	r	j, w	l	lˤ

The VOT of Arabic remains to be fully studied. Academic contributions toward the stops in Arabic are diminishing. The existing literature is scarce, and some of the recent literature is not as promising as the older work that has been conducted. However, one analysis of /b/ and /k/ was carried out by Alotaibi & AlDahri (2011). They instructed 20 subjects to read two words with the two segments (/b/ and /k/), word-medially syllabified as CVCVCV, followed by the low vowel /a/. The average of 20 utterances is 11.35 ms for /b/ and 52.85 ms for /k/. Unfortunately, there is a flaw in this study since some of the participants were non-native speakers of Arabic. Alotaibi and AlDahri's choice to include bilinguals may have contaminated the VOT value, presumably affecting the accuracy of the final results.

Deuchar & Clark (as cited in Khattab, 2000, p. 96) showed that, along the VOT continuum, Arabic voiceless stops occupy the range that is categorized as short lag. A scale that begins from +0 ms and above; yet it does not reach the aspiration range. Voiced stops, on the other hand, have negative voicing (lead voicing) where they get extreme negatives values. That is to say that while Arabic voiceless stops occupy the same value range as that of voiced English stops, Arabic voiced stops and English voiceless stops reside in two unrelated extreme edges (-VOT vs. + VOT). Khattab (2000) concluded in her study that bilinguals who speak Arabic and English produced different VOT for each language. The VOT of monolingual children shifts to the adult-like values as they mature. As far as the bilinguals are concerned, they have built two VOT systems for each language. They also contradicted the earlier assumption by showing a mismatch between the English voiced stops and the Arabic voiceless stops. According to her analysis, Arabic voiceless stops have a wider range for VOT.

A pioneering linguist who investigated the VOT of Arabic was Al-Ani (1970: 31-60). His account of VOT will be used to help provide some background on the dimensions of the Arabic stops.[14] According to his range of duration in ms, we only provide the average. Arabic has a contrast of four sounds in the alveolar place of articulation. In reference to Al-Ani's analysis, the plain voiced alveolar /d/ ranges in the -90 ms[15] and the emphatic voiced alveolar /dˤ/ is -90 ms. The plain voiceless alveolar /t/ is 50 ms and its counterpart emphatic /tˤ/ has the lowest with 25 ms. Moreover, and while the voiceless velar /k/ has the VOT of 70 ms, the voiced bilabial /b/ appears at -85 ms. The last two stops are the voiceless uvular /q/ with a mean of 35 ms, whereas the voiceless glottal /ʔ/ is recorded with the average of 17.5 ms.

Upon a closer look at the dimensions of the stops under Al-Ani's assumptions, the three voiced stops seemed to have high negative values. The alveolar emphatic has been recorded low along the voice-onset time in the positive direction and the voiceless uvular is slightly lower. Still, the glottal stop has the lowest VOT value among all the plosives. It is obvious from the figures that the voiceless sound articulated at the velum reaches the slightly aspirated range as per Cho and Ladefoged's (1999) categorization. Though Al-

[14] Standard Arabic does not have the voiceless bilabial plosive /p/, neither it has the voiced velar plosive /g/. /g/, on the other hand, is found in most varieties of currently spoken Arabic.
[15] Al-Ani does not specify the value direction of the voiced stops. By assuming he implicitly means negativity for the voiced ones, I include the VOT values as negative ones.

Ani mentioned that there were 10 informants used in his study, it is hard to establish the exact number of participants who produced the VOT for each single phoneme.

More recently, Aljutaily and Alharbi (2022) worked on initial stops (/b, d, g, t, k/) of Najdi Arabic (NA) in carrier sentences. The participants of this study were sixty and were divided into two sub-groups on the basis of gender. Just like many other languages, Indo-Europeans and non-Indo-Europeans, the study has shown that speakers of NA produced negative VOT values for voiced stops and positive values for voiceless stops. While the gender had no effect on the VOT durations of voiced stops, it did affect significantly the voiceless counterparts.[16]

The aforementioned studies focused on the phonetics of native Arabic words with no consideration to loanwords. Although there are a few studies that have examined Arabic loanwords in Swahili (Kharusi, 1995; Mwita, 2009; Batais, 2019), none of them looked into the fine details of the adaptation.

2.3 VOT in Swahili

It is essential to comment on the stops of Swahili before talking about its VOT (Swahili consonants are listed in table 3) .[17] As explained earlier in the introduction, the debate between linguists of Swahili has created an uneasy stalemate on the distribution of the plosives along the East Coast of Africa. By referring to some other studies, Kaye & Daniels (1997: 846-847) report that the aspiration is retained in words that have lost nasality from their word-initials. The story is complicated further because aspirated plosives still occur word-medially. They conclude that the aspirated sounds (breathy voice) could be attributed to being “marginal” in the inventory of Swahili stops. Polomé (1967 39-40) has a different consideration of plain/aspirated contrast, for he assigns them the characteristics of fully independent phonemes. He brings this forward because they represent (a) minimal pairs of semantically different lexical items, and (b) contrasts between some noun classes. Implosives, on the other hand, have gotten less fierce disagreement, though this does not mean that they are fully understood. Polomé (1967) again refers to this class of sounds as implosive allophones [ɓ], [ɗ], and [ɠ] of the plosives /b/, /d/, and /ɡ/.

Table 3. Swahili Consonants.

Stops	Fricatives	Affricates	Nasal	Flaps	Glides	Lateral	Implosives
p, pʰ, b t, tʰ, d k, kʰ, ɡ	f, v, (θ), (ð), s, z, ʃ, (ɣ), h	tʃ / tʃʰ	m, n, ɲ, ŋ	r	j, w	l	ɓ, ɗ, ɠ, ʄ

Hayward et al. (1989) investigate Swahili spoken in Mombasa, where they claim that the northern dialect has four distinctions between stops: voiceless aspirated and unaspirated, and voiced pre-nasals and implosives. By comparing the effective maximum contact of alveolars, /t/ is the tensest, followed by /tʰ/ then /ɗ/. As for dentals, they were arranged as /t̪/, /t̪ʰ/, and /ɗ̪/, respectively. To my knowledge the only available study to deal directly with the VOT of Swahili is that of Engstrand & Lodhi (1985). They gave an average of 64.5 ms for aspirated /t/ and 22 ms for the plain /t/. For /k/, they assigned a single word, karo, to two VOT values. The higher value (aspirated) is 66 ms because of its native status, meaning, “washing place.” The lower value (unaspirated) is for the meaning “fee”

[16] This tendency (of positive VOT for voiceless and negative VOT for voiced) is found in American English as in Herd (2020) and Robb et al. (2005).
[17] Swahili consonants in table 3 after Polome´ (1967) cited in (Kharusi 1995: 58). As Polome´ claims, fricatives between parenthesis are from Arabic. Maddieson (1981: 44) adds that implosives are found in the Swahili phonemic system.

with 33 ms due to its being a loan word. Both studies (Hayward et al. 1989; Engstrand & Lodhi 1985) depended on only a single informant, and for this reason, it is difficult to conclude their representativeness of a larger population. In section 3, I discuss the data.

3. Database

3.1 Words

The data design targets the first-place consonants and is limited to stops (plosives). Upon examining the shared stop consonants between Swahili and SA, the data set is narrowed to exclude the inventory of sounds that are foreign to Arabic. Among these Swahili sounds are the voiceless bilabial stop /p/ and the voiced velar stop /ɡ/. Although the latter is a very active allophone in most of the Arabic dialects today, it is not a phoneme of SA. Rather, it is an allophone for the uvular /q/ in many words in the Gulf[18]. The voiced alveolar is also an allophone for the voiced affricate /dʒ/ in Egypt. Therefore, the candidates for the stop position for analysis in this contrastive study are the voiced bilabial /b/, the two alveolars /t/ and /d/, and the voiceless velar /k/. The following task after deciding on the suggested phonemes is to look for Swahili words adapted from Arabic with a stop word-initial.

A list of 35 words was extracted from two Swahili dictionaries (Awde 2011; Kirkeby 2000). Additionally, the glossary of Kharusi's (1995) work was consulted for some insight, though little was adapted from hers. For each stop (/b/, /t/, /d/ and /k/), there is a collection of 7 Swahili words with varying syllables, namely 2 and 3. As we went down the list, we found that loan words in Swahili with the initial /k/ could be an Arabic /k/ or /q/. This forced me to include 7 more Swahili words that feature /q/ initially in the donor language. It is worth mentioning that the 7 words for the stop /t/ in Swahili goes back to two phonemes in Arabic (i.e. 4 to the plain /t/ and 3 to the emphatic /tˤ/). Aside from the Swahili word-initial stop, the Arabic initials /q/ and /tˤ/ will be discussed and contrasted with the Swahili initials to look for alternations. The responses of the Swahili subjects retained a high representation of plosives, so implosives are excluded in the study. With all the words selected for this study, I dedicated some effort toward representative variations in the following vowels across high, mid, and low. It should be noted that the vowels are not meant for any analysis except for some hints if necessary.

3.2 Participants

Two groups were recruited for the analysis part of the study. Each group had a total of 8 subjects. The first was the experimental group. They were 8 males from the city of Mombasa, Kenya. They were born and raised in the city and speak Swahili as a first language. Group two, the control group, had 8 male participants from the city of Buraydah, Saudi Arabia. They were born and raised in the city and speak no language other than Arabic. They can communicate in both Standard Arabic and their regional variety of Arabic. Further, all the informants in the study were instructed to read the list of words prepared for this task in a standardized manner. They were informed to not show any sort of dialectical influence and to say each word at a normal speed. After collecting the audio files, the researcher listened to all the words and measured the VOT duration by using the phonetics software, Praat. A Microsoft Excel spreadsheet was designed for this task by creating columns in which all the measurements were recorded. The total number for all the utterances was 560 and each language had a sum of 280 tokens. After having all the VOT numbers of all speakers, the data frame was loaded into R for statistical analysis. The results of that study are as follows.

3.3 Analysis Procedure

[18] The Gulf countries are Saudi Arabia, Kuwait, Bahrain, Qatar, UAE, and Oman.

VOT measurements were obtained according to Ladefoged & Johnson's (2011: 57-59) description of the waveforms, particularly from the noise burst of the stops until the beginning of the onset of the following vowels. The results were computed to generate R (R Core Team, 2021) by comparing and contrasting all the numbers obtained from the participants. This process was accomplished by examining the contrasts in voicing in the first place, then comparing the segments of the two languages in light of the null hypothesis. The t-test was used to test the null hypothesis; thus, if the null hypothesis was proven correct, then the mean of the two groups would be equal. In addition, the t-test gives no restriction on which sounds or words are in both languages. The test was intended to compare the differences between the two languages, which was based on a sample size of 560 inputs. Now, we proceed to the results.

4. Results and Discussion

In this section, the results of the stops are provided and discussed. First, we would like to investigate the contrast in voicing between the two languages. Second, we comment on the symmetries and asymmetries between the two languages in voiced/voiceless stops.

4.1 Voicing[19]

The distribution of words in the data is evenly divided between the stops. The voiced stops are adapted as their counterparts in Swahili without resorting to the phonetic approximations in the initial stops of the equivalent words. However, the voiceless stops /t/ and /k/ have reflected some mismatch with the source language since there is phonetic approximation of both Arabic /tˤ/ and /q/ to the closest Swahili category (/t/ and /k/). According to the t-test calculations, the data, in general, has presented that there is no significant difference between the voicing categories of loanwords in both languages. This is only an overall observation of the data. It shows that the mean of stops for the entire body of Arabic data is -11.4 ms. Swahili, on the other hand, has a negative value for the mean, even though it is almost in the range of the Arabic one, with -14.0 ms. The negative mean values for the two languages are driven the by the high negative lead found in the VOT of voiced stops. This is clearly observed from the p value of the voiced stops [$t(1.66) = 204$, $p > 0.097$] with a mean of -80.3 ms for Arabic and -74.1 for Swahili. Yet, contrarily, the voiceless adopted the alternative hypothesis as a sequence of rejecting the null one. The t-test presents the p value as significantly different [$t(4.91) = 272$, $p < 0.001$]. The mean value for Arabic is 34.5 ms, whereas it is 26.0 ms for Swahili. For more of a sense on the distribution of stops and their means along the VOT continuum, see figure 1.

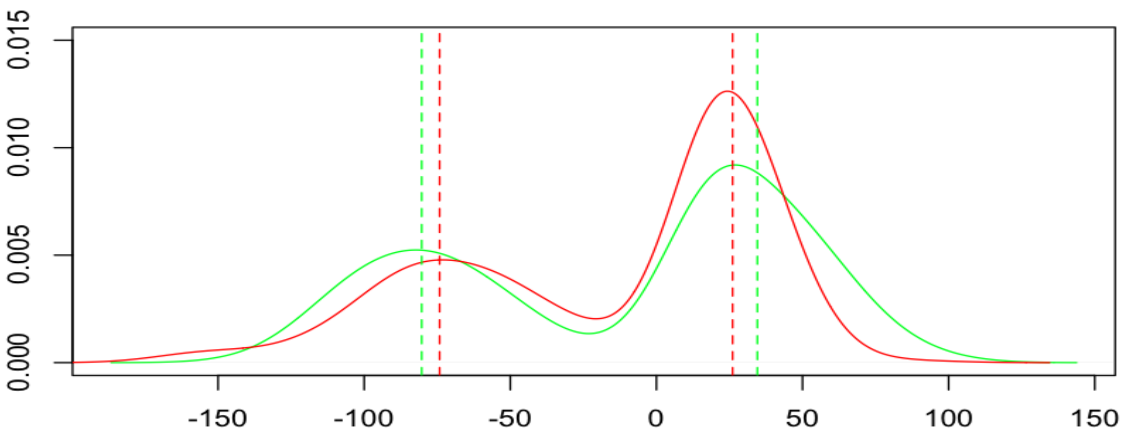

Figure 1. The density plot of Arabic (GREEN) and Swahili (RED) stops [the vertical lines are the means]

[19] The Arabic value comes before the Swahili.

4.2 Voiced Bilabial Stop

The high negative lead of voiced stops has led the two languages to share a common ground on the level of having no substantial asymmetries. The average value for the two segments in Arabic and Swahili is -74.67 and -64.60 ms, respectively, representing quite similar means. Furthermore, the two-sample t-test has assigned voiced alveolar stops a significant value that leads to the acceptance of the null hypothesis. It also gives the p value of [t(-1.96) = 106, p > 0.05]. It should be noted that in this particular data set, the total tokens are 112 divided evenly between the two languages.

4.2 Voiced Alveolar Stop

The comparison between Arabic and Swahili voiced stops reveals that the two languages have a high degree of VOT similarities. This is supported by the close relation between the two languages on the negative side of the graph above (Figure 1). Though there is some slight fluctuation, and the two times show less severe overlaps, in most parts, the two lines appear parallel to each other. Moreover, the means of the two sounds are almost identical, -85.96 for Arabic and -83.76 for Swahili. The increasing of lead voicing around -90 raises the higher value for Arabic. The t-test has given a similar two-sample size and a p value of [t(-0.44) = 96.44, p > 0.65], which accommodates the null hypothesis.

4.3 Voiceless Alveolar Stop

Borrowed words in this category have some variation in initial stops from the origin. We first compare the Arabic /t/ with the Swahili /t/ as a whole. Then we test the Arabic /tˤ/ and its relation to three Swahili words: tabibu, “doctor,” tufani, “flood” and tibu, “scented oil,” for which in all /t/ was initially borrowed from emphatic alveolar of Arabic. After that we compare the Arabic /t/ with the Swahili words that originally adapted from the plain alveolar stop of Arabic only. The null hypothesis for the two /t/s in the two languages of the adapted words in the data shows a significant difference [t(7.85) = 35.96, p < 0.001]. The averages for all the tokens of words beginning with /t/ is 41.65 ms for Arabic and 17.53 ms for Swahili.

The second part of the comparison is between the Arabic VOT of /tˤ/ and Swahili borrowings, which start with /tˤ/ in Arabic (using the three words above). Unlike what the data has just shown in regards to the asymmetry between the /t/s in the two languages, after performing a t-test to test the null hypothesis, the Swahili /t/ that goes back to Arabic /tˤ/ reinforces the hypothesis with a more symmetrical relation between the two segments and maintains more power from the p value [t(0.78) = 45.57, p > 0.43]. The means of the two fall within a very close range with 18.8 ms for /tˤ/ and 17.6 ms for Swahili /t/ adapted from /tˤ/. This result could be motivated by the general representation of Swahili /t/ as lax, which differs from the Arabic light tense[20] /t/. Another piece of evidence for the distinction between Swahili /t/ borrowed from the Arabic initial /t/ comes from the rejection of the null hypothesis of the equal relation. The p value for these two phonemes is statistically and significantly different [t(7.53) = 41.72, p < 0.001].

In our comparison of the Arabic emphatic /tˤ/, statistics have shown that it is closer to the Swahili /t/ than the Arabic plain voiceless stop /t/. However, the basis of the symmetrical relation extends beyond the means. Correspondingly, the p values have given evidence that the difference between Arabic /tˤ/ and Swahili /t/ is insignificant. Further consideration of Swahili /t/, whether inherited from Arabic /t/ or /tˤ/, has also proven no big difference between the Arabic emphatic and the Swahili plain /t/. The p values for the Arabic /tˤ/ (mean = 18.8 ms) vs. the Swahili /t/ from both Arabic /t/ and /tˤ/ (mean = 17.5 ms), and the Swahili /t/ from the Arabic /t/ only (mean = 17.4 ms) are [t(0.906) = 48.71, p > 0.36], and [t(0.788) = 53.87, p > 0.43], respectively.

[20] I hesitate to assign the Arabic voiceless alveolar a tense feature because this contradicts the common literature. However, the mean value of /t/ in Arabic has a higher VOT than its equivalent in Swahili.

4.4 Voiceless Velar Stop

Under the voiceless velar stop, we have six sets to consider. First, we located the difference between Arabic /k/ and Swahili /k/. After that the Swahili /k/ data was divided into two parts. For this step, we set the Swahili /k/ that is inherited from the Arabic /k/ aside from the ones that were transferred from the uvular /q/. Having split the Swahili data, the Arabic /k/ was contrasted with the Swahili /k/ inherited from /k/ and then the /k/ inherited from /q/. Later, we searched for the values of Arabic /q/ in comparison with the Swahili /k/ as a whole, then with the Swahili /k/ inherited from the Arabic /k/ followed by the Swahili /k/ inherited from the Arabic /q/.

For the sake of knowing the values and whether they complement or contradict the null hypothesis, the Arabic /k/ was compared with all Swahili ones (/k/ vs. /k/). The two-sample t-test has shown a very low p value, which led to the rejection of the null hypothesis. The result has given the following output: [$t(11.75) = 99.00$, $p < 0.001$]. While the mean for the Arabic /k/ is 53.5 ms, the average value for the Swahili one is around 30.3 ms. The second test was performed to see if the Arabic /k/ from /k/, not /q/, has any significance. The calculations have proven that they are significantly different as can be noticed from the p value [$t(11.37) = 103.33$, $p < 0.001$]. The means are also drastically different: 53.5 ms for Arabic and 29.5 ms for Swahili. Third, we ran a t-test to see if the Arabic /k/ has equivalence with the Swahili /k/ borrowed from /q/. The p value of this test [$t(9.28) = 99.02$, $p < 0.001$] has led to the rejection of the null hypothesis and presented two extreme means: 53.5 ms vs. 30.4 ms.

As for the uvular /q/, the same task was performed in order to figure out whether the null hypothesis for the VOT of Arabic /q/ equals the Swahili /k/, /k/ from /k/ or /k/ from /q/, respectively. It has been found that /q/ and /k/ are significantly different since the means are varying: 18.2 vs. 30.3 with a p value [$t(-10.41) =152.1$, $p < 0.001$]. A test was conducted as well to test the relationship between /q/ and the Swahili /k/ from /k/. Again, the alternative hypothesis was accepted and a significant difference has been found [$t(-8.14)= 71.49$, $p < 0.001$], with a mean of 18.2 ms vs. 29.5 ms. Furthermore, the comparison between /q/ and the Swahili initial form /q/ has yielded a rejection of the null hypothesis because of the small p value [$t(-6.40) = 54.10$, $p < 0.001$]; the means of 18.2 ms and 30.4 highlight a considerable gap between the two values.

After having considered the results and the variations between the two languages' VOTs, we proceed to the conclusion.

5. Conclusion

Many of the topics this article has discussed have been at the center of the debate between the linguists who are interested in the status of stops in Swahili. They vary on their judgments, yet they all agree that the language has aspiration. To what extent the aspiration is actively implemented among the speakers of the two countries (Kenya and Tanzania) is still unclear. One scholar theorizes that the feature of aspiration is found more often in the North than in the South (see Polomé, 1967). Another goes on to claim that aspiration is mostly dependent on the parts of speech. Not all speech categories are targeted, but he defined it as a certain type of noun-class (as in Engstrand & Lodhi, 1985). The disagreement is quite intense, but no clear-cut evidence on the final production exists.

It has been found in some of the literature reviewed that aspiration is completely lost in the loanwords from Arabic except for a few isolated cases. The strange thing about these arguments is that the implosives are minimally discussed. Kharusi (1995) and Polomé (1967) claimed that the implosives are allophones of the voiced phonemes of the stops. Furthermore, Kharusi believed that the adaptation has enforced the preservation of the implosive category through phonetic adaptation. This is may be due to the extensive usage of implosives over voiced plosives. In the data, a kind of implosive voicing can be observed, though I hesitate to generalize the results. The fluctuation stems from the variation in the products of the subjects. The waveform of the initial voiced segment fails to keep the consistency of increasing voicing spikes as it approaches the vowel onset.

Cross-linguistically, languages differ in their VOT dimensions (see table 1 above). Lisker and Abramson's (1964) analysis gave the voiced stops negative values except for English. In their data, all the voiceless classes were positively assigned VOT numbers. Lisker and Abramson's work has also motivated other linguists to investigate their own languages' VOTs. Unlike English, Arabic's stops have been measured and studied. Although Al-Ani's (1970) work is over 46 years old, it is still the most detailed and valuable work on the VOT of Arabic. Nevertheless, his assumptions lack complete accuracy since the number of the subjects for the measurements in his study is unknown. Swahili, on the other hand, desperately lacks any comprehensive consideration of its stops' VOTs. Engstrand & Lodhi's (1985) study could be considered the most pioneering of all Swahili studies on VOT, yet it dedicates only one section to the topic of VOT, and the findings in this section were extracted from merely one speaker living abroad. This casts doubt on the ability of their results to be statistically generalized to larger populations.

The result of the study (table 4) has revealed that the adaptation of the VOT is a combination of phonological and phonetic approximation; none of them has succeeded over the other. The phonetic adaptations are very clear from the voiced stops. The two segments /b/ and /d/ in Swahili have shown means closer to their counterparts in Arabic. The p value of the two voiced segments proves that their differences with Arabic are insignificant. Voiceless initial stops of Swahili are in favor of the phonetic adaptation. Though they present the same Arabic phoneme, however, their VOTs are lower than in Arabic. They have displayed significant differences from Arabic, which consequently led to the acceptance of the alternative hypotheses. The only exceptions from the voiceless class are words that feature the initial /t/, from the Arabic /tˤ/. The means of these two segments in both languages have phonetically accepted the null hypothesis and rejected the alternative one.

Table 4. A summary of VOT of stops in adapted loanwords from Arabic.

Consonant	Arabic VOT Mean	Swahili VOT Mean
/b/	-74.6	-64.5
/d/	-85.9	-83.7
/t/ words from Arabic /tˤ/	18.8	17.6
/t/ words from Arabic /t/	41.6	17.4
/k/ words from Arabic /k/	53.5	29.5
/k/ words from Arabic /q/	18.2	30.4

For further investigation, a VOT measurement for the phoneme /p/ in Swahili is tested for aspiration/non-aspiration contrast. A total of 6 native Swahili words (puto 'ballon',punda 'donkey' ,pombe 'alcohol', pembetatu 'triangle', panya 'mouse' ,pamba 'cotton') and 1 loan word (pilipili 'pepper') were measured in Praat and analyzed in R. The mean for all 56 tokens by 8 speakers was 19.67 ms. The mean for the only Arabic word in the list as elicited by 8 different speakers was 21.31 ms. By comparing the Arabic loan with the Swahili words, the t-test has shown that there is no significant difference between the two measurements.

In conclusion, we wish to emphasize that the results are meant to compare and contrast the dimensions of the VOT of the Arabic words in Swahili for stops in the initial position. As we have seen from /p/ above, the aspiration is not tested in Swahili. That would posit a question: Is aspiration is really contrastive in the language? Future studies should contrastively investigate all the stops of the two languages and include both loan and native dimensions to see if the results are still consistent. Factors such as location, social status, ethnicity, and religion, if considered, could have added more valuable results to the data.[21]

Acknowledgements

Researchers would like to thank the Deanship of Scientific Research, Qassim University for funding publication of this project

References

[1] Al-Ani, S. (1970). *Arabic phonology; an acoustical and physiological investigation*. The Hague: Mouton.

[2] Aljutaily, M., & Alharbi, B. (2022). Voice Onset Time of Initial Stops in Najdi Arabic. *World Journal of English Language, 12*(6) 83-96.

[3] Alotaibi, Y. A., & AlDahri, S. S. (2011). Investigating VOTs of Arabic stops/b, k/with comparisons to other languages. *Image and Signal Processing (CISP), 4th International Congress, 5*, 2413-2417.

[4] Alsamaani, M. (2021). *VOT and F0 in the production and perception of Swahili obstruents: from the island to the coast to the inland region*. Doctoral dissertation: University of Calgary.

[5] Awde, N. (2011). Swahili practical dictionary. New York, NY: Hippocrene Book, INC.

[6] Batais, S. (2013). *Consonantal and syllabic repairs of Arabic and Dutch loanwords in Indonesian: A phonological account*. Gainesville, FL: MIT dissertation.

[7] Batais, S. (2019). Consonantal Adaptation of Arabic Loanwords in Kiswahili and Kisukuma: A Phonological Explanation. *Advances in Language and Literary Studies, 10*(2), 85-94.

[8] Bin-Muqbil, M. S. (2006). *Phonetic and phonological aspects of Arabic emphatics and gutturals*. Doctoral dissertation: University of Wisconsin-Madison.

[9] Chao, K. Y., & Chen, L. M. (2008). A cross-linguistic study of voice onset time in stop consonant productions. *Computational Linguistics and Chinese Language Processing*, *13*(2), 215-232.

[10] Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, *27*(2), 207-229.

[11] Engstrand, O., & Lodhi, A. Y. (1985). On aspiration in Swahili: Hypotheses, field observations, and an instrumental analysis. *Phonetica*, *42*(4), 175-187.

[12] Frankl, P. J. L. (1991). The nature of aspiration in Swahili. *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, *141*(2), 366-375.

[13] Hayward, K. M., Omar, Y. A., & Goesche, M. (1989). Dental and alveolar stops in Kimvita Swahili: An electropalatographic study. *African Languages and Cultures*, *2*(1), 51-72.

[14] Herd, W. (2020). Sociophonetic voice onset time variation in Mississippi English. *The Journal of the Acoustical Society of America, 147*(1), 596–605.

[15] Ito, C. & Kenstowicz M. (2013). The Adaptation of contemporary Japanese loanwords in Korean. *Japanese/Korean Linguistics, 22*, 3-20.

[16] Kang, Y. (2010). The emergence of phonological adaptation from phonetic adaptation: English loanwords in Korean. *Phonology*, *27*(02), 225-253.

[17] Kaye, A. S., & Daniels, P. T. (Eds.). (1997). *Phonologies of Asia and Africa: Including the Caucasus*. Winona Lake, IN, USA: Eisenbrauns.

[18] Kharusi, N. S. (1995). *The linguistic analysis of Arabic loan-words in Swahili*. Doctoral Dissertation: Georgetown University.

[19] Khattab, G. (2000). VOT production in English and Arabic bilingual and monolingual children. Nelson, D. & P. Foulkes (eds). *Leeds Working Papers in Linguistics 8,* 95-122.

[21] For a recent more-detailed study on Swahili, see Alsamaani (2021).

[20] Kirkeby, W. (2000). *English Swahili Dictionary*. Skein: Norway.

[21] Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, *18*(4), 686-706.

[22] Ladefoged, P., & Johnson, K. (2011). *A Course in phonetics*. Wadsworth, Cengage Learning.

[23] Lisker, L., & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384-422.

[24] Lodhi, A. Y. (2003). Aspiration in Swahili adjectives and verbs. *Africa & Asia*, *2*, 155-160.

[25] Maddieson, I. (1981) *UPSID: UCLA Phonological Segment Inventory Database (UCLA WPP 53).* Los Angeles, CA: University of California.

[26] Mwita. L. C. (2009). The adaptation of Swahili loanwords from Arabic: A constraint-based analysis. *The Journal of Pan African Studies 2*(8), 46–61.

[27] Nakai, S., & Scobbie, J. (2016). The VOT category boundary in word-initial stops: Counter-evidence against rate normalization in English spontaneous speech. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *7*(1), 1-31.

[28] Netelenbos, N., Li, F., & Rosen, N. (2016). Stop consonant production of French immersion students in Western Canada: A study of voice onset time. *International Journal of Bilingualism, 20*(3), 346– 357.

[29] Paradis, C., & LaCharité, D. (2011). Loanword adaptation: From lessons learned to findings. *Handbook of Phonological Theory,* (2), 751-778.

[30] Polomé, E. C. (1967). *Swahili language handbook.* Washington, D.C.: Center for Applied Linguistics.

[31] Port, R. F., & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *The Journal of the Acoustical Society of America*, *66*(3), 654-662.

[32] R Core Team (2021). R: A language and environment for statistical computing (Version version 4.0.2) [Computer software]. R Foundation for Statistical Computing. URL https://www.R-project.org/.

[33] Robb, M., Gilbert, H., & Lerman, J. (2005). Influence of gender and environmental setting on voice onset time. *Folia Phoniatrica et Logopaedica, 57*(3), 125–133.

[34] Trask, R. L. (2007). *Trask's historical lingistics*. Revised by Robert McColl Millar. London: Hodder.

[35] Tucker, A. N., & Ashton, E. O. (1942). Swahili phonetics. *African studies*, *1*(2), 77-103.

[36] Velupillai, V. (2012). *An introduction to linguistic typology*. John Benjamins Publishing.