A PHONETIC STUDY OF BREATHY VOICING IN DZA

by

OTHANIEL, NLABEPHEE KEFAS

SUPERVISORS: Ioana Chitoran

Mark Van de Velde

A THESIS SUBMITTED TO UNIVERSITÉ PARIS CITÉ IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER'S DEGREE IN LANGUAGE SCIENCES IN PHONETICS AND PHONOLOGY, PARIS GRADUATE SCHOOL OF LINGUISTICS

AUGUST, 2022

Acknowledgement

I appreciate Université Paris Cité for offering me the MIEM scholarship of excellence. I want to appreciate Ian Maddieson for giving me audience and for his valuable comments which helped in structuring the Dza breathy versus modal vowels questionnaire. I thank Dmitry Idiatov and Mark Van de Velde for their so many supports in this study and in other previous studies. I thank Jalaal Al-Tamimi for unreserved help rendered to me in organizing my data, allowing me to use his Praat script and modifying the script to suit the need of my project, and for introducing me to VoiceSauce. My special appreciation goes to Professor Ioana Chitoran for never giving up on what I was doing.

Table of Contents

Table of Figures

List of Tables

Abbreviations

1	First person
GEN	Genitive marker
GRN	Gospel Recording Network
HAB	Habitual
LOC	Locative
NEG	Negation
NPST	Non Past Tense
PFV	Perfective
POS	Possessive
2	Second person
SG	Singular
SPEC	Specifier
3	Third person

1 Introduction

Voicing is the production of sound primarily initiated in the larynx by the vibration of the vocal folds. We can differentiate people by the sound of their voice, and we, humans, can even attempt to imitate how other people speak by manipulating the quality of our regular voice to sound like the people we are imitating. One of the cardinal components that gives us the leverage to do so as a specie is our ability to manipulate the larynx through momentary modification of the frequency at which our vocal folds vibrate and by adjusting the degrees of constriction of relevant tissues in the larynx (Laver 1980; Garellek 2019; Esling et al 2019). Studies have shown that demography, emotions, physiological wellbeing, and many other factors could make people's voice vary (see Klatt and Klatt 1990; Kreiman and Gerrat 2002; Suire et al 2020) However, there are variation of voice that are specific features of languages shared among speakers. In spoken languages, voice quality variation plays a major role in describing the speech material of individual languages. Perceptually, this kind of variation is employed by speakers for the purpose of linguistic and paralinguistic communication (Laver 1980, Esling et al 2019). In languages where voice quality variation is relevant for linguistic communication, speakers actively contrast between segments by phonation types, viz., voiceless, breathy, modal, creaky voice etc (Ladefoged and Maddieson 1996; Gordon and Ladefoged 2001). This study explores the contrast between the modal and the breathy phonation in Dza, a language spoken in the Benue – valley region of North – Eastern part of Nigeria. The languages of this region are to a large extent underdescribed, so little is known about the phonetic structures of these languages. The research is based on my earlier auditory impressionistic analysis of the phonetics and phonology of the Dza language. The goal of this current project is firstly to investigate the acoustic correlates of breathy and modal vowels in Dza, and secondly to investigate to what extent is the modal – breathy contrast systematic in Dza.

The work is organized into eight sections. This first section gives a general introduction to the research. Section 2 contains a literature review on modal versus non-modal phonation. Section 3 contains sociolinguistic information on the Dza language and aspects of its phonetics and phonology from previous research. Section 4 gives the hypothesis and the predictions of the study. Section 5 is a description of the methodological approach of the research. The results are presented in section 6 while section 7 is the interpretation of the results, and section 8 is the conclusion and recommendations for further studies.

2 Modal versus non-modal voicing in vowels

2.1 Articulatory setting and phonetic properties of modal and non-modal voicing

The vowel systems of the world's languages are described in terms of features, and contrastive voice quality is one of those features that distinguish vowels apart in some languages (Ladefoged and Maddieson 1996). In terms of laryngeal articulation, this feature is presented in a continuum ranging from a glottal state in which the vocal cords are in a voiceless position with the arytenoid cartilages far apart, to a state involving a creaky voice with the arytenoid cartilages firmly pressed together (Gordon and Ladefoged 2001; Garellek 2009). This description suggests a glottal state somewhere in the mid-region of the continuum that is characterized by a regular vibration of the vocal cords and the arytenoid cartilages are close together but not firmly in a medial posture. It is at this neutral laryngeal setting that modal voiced vowels are produced. Thus, modal voicing is defined as the neutral mode of phonation that is characterized by a regular vibration of the vocal folds with absence of audible friction, and a moderate approximation of the arytenoid cartilages (Esling et al 2019; Laver 1980; Ladefoged and Maddieson 1996). By implication, any phonation that is produced at other laryngeal settings along the continuum described in Figure 1is a non-modal phonation. Figure 1 gives a simple illustration of the phonation type continuum:

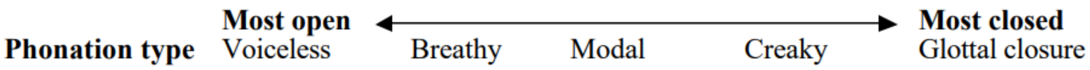


Figure 1: Continuum of Phonation types (Gordon and Ladefoged 2001)

There are sub-classifications of the various phonation types illustrated in Figure 1 (For detailed description of the laryngeal settings of the various phonation types consult Laver 1980, Esling et al 2019, Ladefoged and Maddieson 1996). However, the phonation types breathy, modal, and creaky phonation are considered to be the most common phonation types that are used in languages (Garellek 2019; Gordon and Ladefoged 2001). Further discussions will focus on these three main phonation types. As such the term ‘breathy phonation’ or ‘creaky’ phonation does not connote an absolute point on the continuum, rather they each cover range of settings. For example, on the breathy side of the continuum some sounds are breathier than the others, similarly the creaky phonation continuum covers a range of settings with vocal fry at the extreme (Garellek 2019).

These three phonation types are described relative to each other in various voice quality studies (See Garellek 2019; Gordon 1998; Gordon and Ladefoged 2001, Kreiman et al 2014). When compared with modal phonation as described above, the breathy phonation is characterised by a slightly opened glottis with higher rate of airflow, the arytenoid cartilages are apart such that the vocal codes are vibrating but without substantial contact (Ladefoged and Maddieson 1996). Creaky phonation on one side is characterised by a lower rate of airflow than modal phonation, the arytenoid cartilages are pressed together but not too tight to restrict voicing, such that only a small length of ligamental vocal cords is vibrating (Ladefoged and Maddieson 1996). In terms of acoustic properties, modal phonation is associated with regular periodicity void of controllable noise, substantial overall acoustic intensity, medium range of fundamental frequencies, intermediate values of spectral tilt when compared to non-modal phonation types (Gordon and Ladefoged 2001; Laver 1980). In contrast, breathy phonation is associated with an increase in spectral noise at higher frequencies, a decrease in overall acoustic intensity, a lower fundamental frequency and a negative spectral tilt compared to modal phonation. Creaky phonation, on the other hand, is characterized by aperiodic glottal pulses, a decrease in overall acoustic intensity, a strongly positive spectral tilt, and a lower fundamental frequency (Gordon and Ladefoged 2001; Garellek 2019). Gordon and Ladefoged (2001) noted, however, that the lowering of the fundamental frequency is not universal. Similarly, formant frequencies, duration and aerodynamic properties show a phonation contrast in some languages and in other languages this is not the case. This is because several different laryngeal actions can be used to achieve the perceptual effect of breathiness or creakiness (Blankenship (2002).

2.2 Acoustic parameters for differentiating phonation types

Various studies use different sets of parameters and methods to measure voice quality. Here is a representative list of parameters:

- H1 – A1 - the difference in amplitude between the first harmonic and the most prominent harmonic in the F1 (Gordon and Ladefoged 2001; Wayland and Jongman 2002; Keating et al 2021)
- H1 – A3 - the difference in amplitude between the first harmonic and the most prominent harmonic in the F3 regions (Gordon and Ladefoged 2001; Wayland and Jongman 2002; Keating et al 2021)
- H1 – H2 – the difference in amplitude between the first and second harmonics (Blankenship 2002; Gordon and Ladefoged 2001; Garellek 2019; Keating et al 2021; Wayland and Jongman 2003)
- H2 – H4 - the difference in amplitude between the second and fourth harmonics (Kreiman et al. 2007; Keating et al 2021; Garellek 2019)
- H4–H2kHz - the difference in amplitude between the fourth harmonic and the harmonic closest to 2000Hz, (Keating et al 2021; Garellek 2019)
- H2kHz–H5kHz - the difference in amplitude between the harmonic closest to 2000 Hz and the one closest to 5000 Hz (Keating et al 2021; Garellek 2019)
- HNR – harmonics to noise ratio (Keating et al 2021; Garellek 2019)
- CPP – cepstral peak prominence (Garellek 2019; Blankenship 2002)
- Fundamental frequency (Wayland and Jongman 2002; Gordon and Ladefoged 2001; Garellek 2019)
- RMS energy (Wayland and Jongman 2002; Esposito and Khan 2012)

The first six parameters are measures of spectral tilt. Spectral tilt correlates with the laryngeal articulation of spreading/constriction of the vocal fold. An increase in spectral tilt corresponds to vocal fold spreading, while a decrease in spectral tilt corresponds to constriction (Garellek 2019, Blankenship 2002). Since breathy phonation is associated with vocal fold spreading, spectral tilt measurements of breathy vowels should be higher than those of modal vowels. Similarly, modal vowels should have higher spectral tilt measures than creaky vowels since creaky phonation is associated with constriction. As indicated in the list, some studies use spectral tilt measures over different frequency bands (H1 - A1, H1 - A3), others use harmonic-based measures, while others use both. Garellek (2019) states that the two methods of measuring spectral tilt are correlated with each other and may even overlap. For example, in the psychoacoustic model of voice devised by Kreiman et al 2004, the measures of H1-H2, H2-H4, H4-H2k, H2k-H5k are taken together as sufficient measures that characterize spectral tilt in various harmonics and frequency bands (Garellek 2019; Keating et al 2021).

CPP and HNR are measures of spectral noise. When “the spectrum of a periodic signal shows well-defined harmonics; its cepstrum has a prominent peak at a quefrency corresponding to the duration of the f0 cycle” (Blankenship 2002). In 2.1 above, I mentioned that modal phonation is characterized by regular periodicity when compared to non-modal phonation types. Thus, modal vowels are expected to have higher cepstral peaks in comparison with non-modal vowels since the non-modal phonation types are associated with less periodicity. Similarly, the aperiodic noise associated with the non-modal phonation types will lower the HNR when compared to modal phonation that has regular periodicity (Blankenship 2002; Garellek 2019; Klatt and Klatt 1990; Gordon and Ladefoged 2001).

F0 and RMS energy are measures of vibration frequency and its amplitude. The less tensed state of the vocal folds during breathy phonation associates it consistently with lowered F0 relative to modal phonation (Klatt and Klatt 1990; Gordon and Ladefoged 2001; Blankenship 2002; Garellek 2019). Similarly, creaky phonation is also associated with lowered F0. However, Gordon and Ladefoged (2001) states that the lowering effect in creaky phonation is not universal, certain languages have developed high tone as a reflex of glottal constriction. Both breathy and creaky phonation types are associated with decrease in intensity when compared with modal phonation. Decrease in intensity correlates with decrease in overall amplitude measured by RMS energy (Gordon 1998; Gordon and Ladefoged 2001).

2.3 Distribution of non-modal vowels

Most languages of the world do not have contrastive phonation types in vowels. As such, one could say modal vowels are more common in the languages of the world. Gordon (1998) proposed that the rarity of non-modal vowels across languages has a perceptual basis. For example, in tonal languages non-modal phonation reduces the ability of vowels to manifest tonal contrasts in a salient way, so that these languages limit the overlap between tonal and phonatory contrasts (Gordon 1998). Other acoustic properties that make non-modal vowels less salient than modal vowels include:

- Decrease in overall acoustic intensity resulting in a decrease in loudness,
- The disruption of formant structure due to the reduced intensity of non-modal phonation which, he argues, makes the recovery of vowel contrasts more difficult,
- The reduced effect of non-modal phonation on vowel duration. Gordon cited the examples of Kedang and Jalapa Mazatec where non-modal vowels occur as phonetically longer segments, up to 50% longer than their modal counterparts. However, non-modal phonation is limited to the first half of the vowel, while the second half of non-modal vowels is characterized by modal voicing, etc.

Contrastive non-modal vowel phonation spread across various language families in the world including languages of sub-Saharan Africa (Maddieson 2018; Keating et al 2021).

2.4 History of non-modal phonation study in Dza

In Othaniel 2016, I reported that Dza employs three phonation types for sound contrast: voiceless, modal, and breathy phonation. In the phoneme inventory of Dza, I illustrated a nearly symmetrical modal voiced and voiceless consonant contrast across different places and manners of articulation. For the vowels, I illustrated that Dza has nine (9) modal vowels /i, e, ɛ, ɨ, ə, a, u, o, ɔ/. I showed that the modal vowels /i, ɛ, ɨ, ə, a, u, ɔ/ can be articulated as either oral or nasal vowels and that the contrast is phonemic. Then, I stated that the Dza modal vowels /i, e, ə, a, u, o/ are contrastive with another set of non-modal vowels which I perceived as breathy voiced vowels /i̤, e̤, ə̤, a̤, ṳ, o̤/. However, I interpreted the modal – breathy contrast to be a surface phonetic feature. My interpretation was based on three apparent restrictions I observed in the distribution of breathy vowels in Dza. First, the breathy vowels in Dza carry only non–high tone; secondly, they only occur after voiceless consonants and thirdly, I had an intuition that the modal – breathy contrast is utilized only when speakers call words in isolation, but the contrast is lost during speech. I proposed two initial hypotheses from the observed restrictions: (1) What I perceived as breathy voicing on vowels is a rudiment of contrast between aspirated and non-aspirated consonants. (2). Thanks to Mark Van de Velde and Dmitry Idiatov (personal conversation), my attention was drawn to the fact that the so-called breathy segments seem to have a lower pitch when compared to the three level tones on modal segments in Dza. Thus, what I called modal – breathy contrast may instead be a case of tonogenesis, denoting an emergence of a fourth level tone in Dza.

A third hypothesis came forth from my phonological comparative study of the Bikwin – Jen languages. I reported that breathy vowels occur in all the languages of the Bikwin – Jen cluster except for Moo and Kyak (See Othaniel 2017). In addition to my claim, Harley (2020) also reported that he came across a breathy segment in Kyak. However, I noted that the breathy vowels are larger in number in Dza when compared to other Bikwin – Jen languages. In Norton and Othaniel (2020), we showed that the breathy feature repeatedly co-occurs with reduction of obstruents, viz., devoicing, debuccalisation of the proto-Bikwin – Jen voiceless fricative *f in Dza, and development of the voiceless labial-palatal [ɥ̊] in Munga Doso and Dza. Thus, we surmised that the breathy voice is a trace of consonant change in Dza. Furthermore, we directed a question that suggests breathy phonation in Dza might be a feature of the entire root instead of just vocalic segments.

Table 1: Breathy vowels after obstruents in Bikwin - Jen languages (Norton and Othaniel 2020)

gloss	Burak	Loo	Maghdi	Mak	Kyak	Moo	Leelau	Tha	Doso	Dza
‘leg’	*vō̤*	*vō̤*	*vò*	*bō*	*bō*	*bō*	*bò*	*bò*	*bù*	*pū*
‘fencing mat’					*gē*		*gəì*	*gī̤*	*gì̤*	*kī̤*
‘answer (v.)’	*gàb*	*gə̀b*	*gàb*	*gə̀b*	*gwə̀b*	*gwə̀b*	*gwə̀b*	*gwə̀*	*gwə́ŋ*	*kwə̤̀*
‘neck’	*dòl*	*dūl*	*dùl*	*dūl*	*dūl*	*dúl*	*dwùl*	*ɟwí*	*d͡ʒwìm*	*t͡ʃwī̤*
‘stand’	*dē̤*	*dē̤*	*də̤ì*	*də̤ì*	*dɛ̄*	*dɛ̄*	*dē*		*d͡ʒì*	*t͡sɨ̀*
‘snake’		*bjə̀k*	*bjə̀k*	*bjə̀k*	*bjɔ́k*	*bjə̀ŋ*	*bjēk*	*ɟəù*	*d͡zəù*	*t͡sō̤*
‘feather’					*bjə̀ŋ*	*bjə̄ŋ*	*bjə̀ŋ*	*ɟɛ̂*	*d͡zàm*	*t͡sà̤*
‘say’	*gē̤*	*gē*	*d͡ʒē̤*	*gəì*	*d͡ʒē*	*d͡ʒē*	*gəì*	*kìjè*	*kɨ̀ŋ-kâ*	*t͡sà*
‘elephant’	*d͡ʒʊ́k*	*d͡ʒʊ́k*	*d͡ʒɔ̀k*	*zʊ́k*	*zòk*	*zɔ̀k*	*zɔ̀k*	*ðəù*	*zò*	*sō̤*
‘count’					*zēn*	*zēn*	*ðə̂n*	*hò*	*zàm*	*sà̤*
‘ashes, dust’					*vùk*				*vəù*	*fō̤*
‘new’	*fṳ̄*	*fŭ*	*fù*	*fū:*	*fú*	*fú*	*fú*	*fù*	*fû*	*hú*
‘blood’	*wě*		*ɥē̤*	*ɥē*	*zwī*	*zǐ*	*zī̄*	*vì*	*ŋ̩̀ɥ̊ì̤*	*ɥ̊ǐ̤*
‘sand$_B$’	*d͡ʒwē̤*	*ʒwē*	*ʒwəì*	*zwē*	*zwī*	*zí*	*zwí*	*vī*	*ɥ̊ì̤*	*ɥ̊ī*
‘lion’		*d͡ʒwē̤*	*d͡ʒwȁ*	*zwà̤*	*zwā̤*	*zwà*	*zwà*	*zwà*	*ɥ̊à̤:*	*ɥ̊ā̤*
‘smoke’	*wṳ̄*	*dùŋɔ̀*	*wú lwà*	*dû*	*d͡ʒù*	*d͡ʒū*	*dù*	*ðɨ́ú*	*ɥ̊ì*	*ɥ̊ī*

In the first group of roots in Table 1, we have given examples where Dza has voiceless onsets followed by a breathy vowel, the related lexical items in the other Bikwin - Jen languages have voiced onsets. The last set of cognates in this group for the word 'new' shows an example of

debuccalization where the proto Bikwin – Jen labial fricative /*f /is weakened to a voiceless glottal fricative /h/.
Furthermore, Harley (2021) reported that he came across a word in Kyak with the laryngeal feature of breathy phonation. It was the Kyak word [ǹ̤n̤ɛ̤̀] ‘four’ which he compared with [nɛ̂] ‘mother’. Figure 2 shows waveforms and spectrograms of these words from Harley 2020:

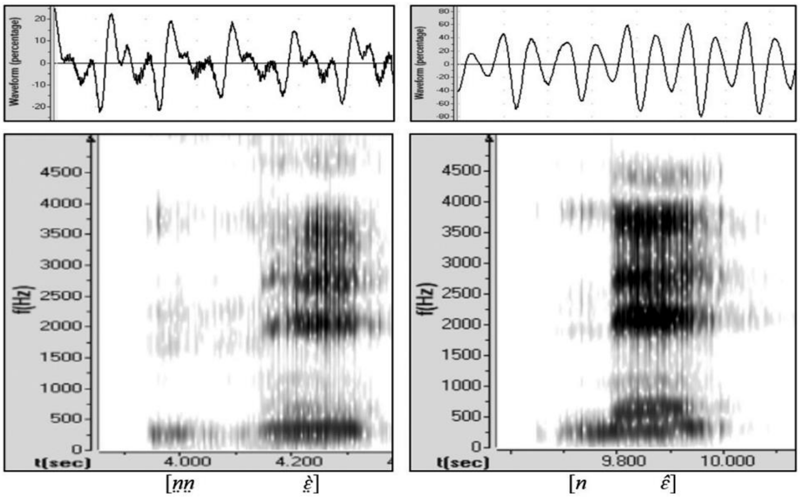
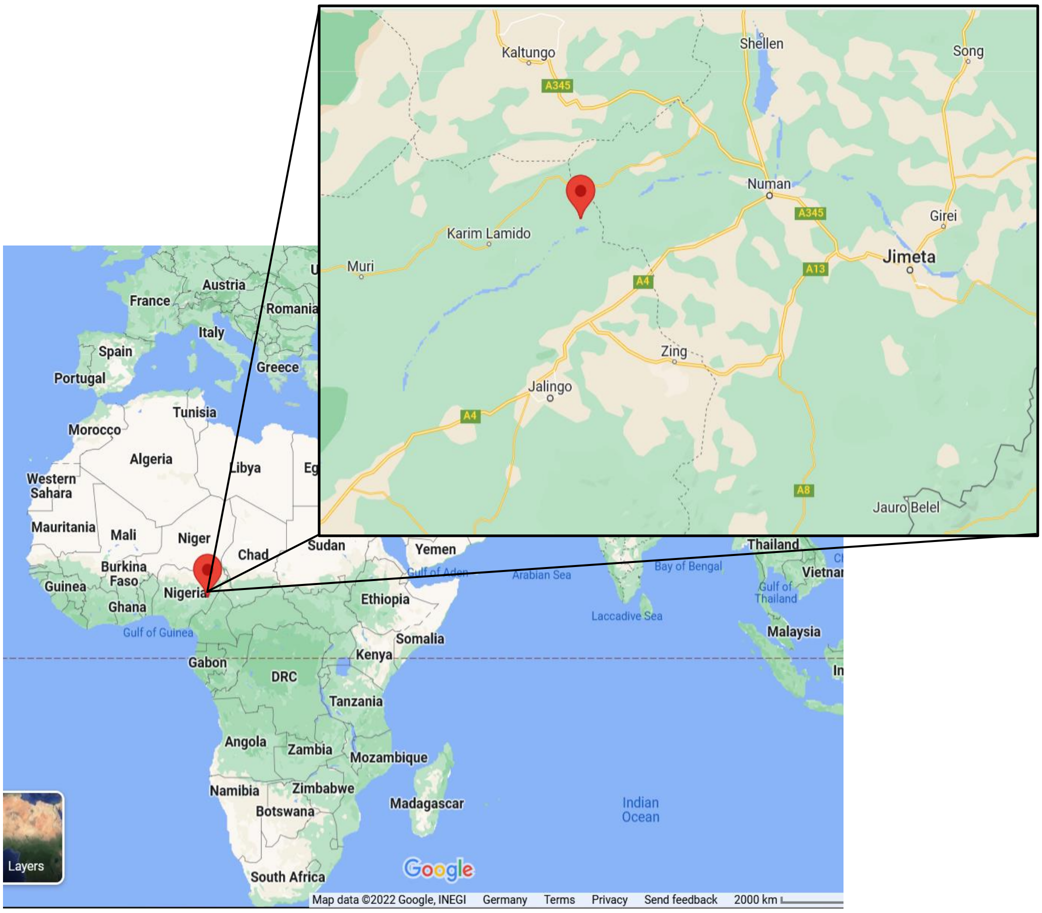


Figure 2: Waveforms and spectrograms of breathy and modal voiced segments in the Kyak words [ǹ̤nɛ̀] 'four' and [nɛ̂] ‘mother’ in Harley (2020)

From the visual cues, the segment on the left shows some characteristic of breathy phonation as outlined by Gordon and Ladefoged (2001). The left segment waveform shows significant aperiodic noise compared to its right counterpart, and the left segment spectrogram shows a decrease in intensity, particularly at high frequencies, compared to its right counterpart. However, the complex nasal onset in the left segment makes the interpretation of the visual cues from the waveform display and spectrograms in Figure 2 somewhat ambiguous. We cannot rule out a possible spread of hypernasality from the complex nasal onset to the vowel since breathy phonation and nasality are acoustically similar (see Garellek et al 2016).
The foregoing discussion initiated my interest to investigate: (1) the correlation between breathy voiced vowels and non–high tones in Dza; (2) the relatedness of breathy vowels and the consonant segments that precede them in words. In order to answer the above questions, it is necessary to carry out a systematic instrumental phonetic analysis of the modal – breathy contrast in Dza vowels, to determine the acoustic characteristics of voice quality difference between the two sets of vowels.

3 The Dza language

Dza is a minority language spoken in the North-Eastern part of Nigeria. The speakers are concentrated around Jen town on the upper bank of the river Benue at the border between Gombe, Adamawa and Taraba state. This research is based on data collected from speakers of Dza who are born and brought up in Jen town Karim-Lamido local government, Taraba State. Jen is located at the north bank of the Benue River, north-eastward from Lau, west of Numan town around the geographical coordinates 9°22'46.5"N 11°27'55.4"E. Figure 1 below is a map showing the geographical location of Jen.



3.1 Linguistic Classification

Dza is part of a cluster of ten languages known as the Bikwin – Jen. The cluster makes part of a sub-group of the Niger – Congo family called the Adamawa languages. Adamawa is a cover term which is currently used for a number of small and underdescribed linguistic groups in this part of Africa, whose affiliation to the Niger-Congo phylum is generally beyond doubt but whose exact position within this phylum remains unclear (see Bennett and Sterk 1977 for comments). The group was originally put together by Joseph Greenberg. Dza and Munga-Dosso made up Greenberg's group 9 of the former Adamawa-Eastern group of Niger–Congo called 'Jen' (Greenberg 9). After Greenberg's work, the genealogical unity of the new Adamawa-Eastern group and its classification as Niger – Congo was debated (see Samarin 1971 for example). Other studies proposed that Dza and related languages of the Benue valley in the 'adamawa' group are related to the Gur languages (see Bennett and Sterk 1977; Kleinewillinghöfer 1996). More recently, Dmitry Idiatov and Mark Van de Velde (2020) claim that these languages belong rather to the Benue – Congo family. At the moment, major reference materials vary the positioning of the Bikwin – Jen languages in the Niger – Congo genealogy. In the updated version of glottolog, Dza and sister Bikwin – Jen languages are classified in a group called Cameroon – Ubangian (Hammarström et al 2022). This group appears earlier in the Niger – Congo genealogy when compared to 'Adamawa – Ubangi' subgroup in the ethnologue classification (see Ebarhard et al 2022). However, one of the significances of this research is that it will contribute to the ongoing efforts for the internal reconstruction of the Benue – Valley languages especially in the AdaGram research project of LLACAN.

3.2 Dialects

Dza town is divided into three regions: Dzakə, Ye and Nwabang. Historically, each of these regions speak a different dialect of the language. But this is not the case with the younger generation of Dza speakers. As member of the community who spent time in Dza, I daresay Dza speakers below age forty-five (45) speak a form of the language that is a merger of the three dialects. The rudiments of the dialectal differences are in the speech of speakers above age 45. The three regions are represented in the data I collected from previous fieldworks in the Dza area. The dialectal differences I observed include voicing alternation of consonant segments, palatalization of the voiceless alveolar sibilant [s] in the Nwabang dialect in comparation with Dzakə and Ye, fortition of the voiced alveolar sibilant [z] into affricates [dz] in [tʃ] in Nwabang and Dzakə, and syntactic difference in the internal structure of the noun phrase [g͡bə́kə́dàŋ]. Table 2 below shows the lexical difference observed from my previous word collection exercises:

Table 2:Dialectal Differences in Dza

Word	**Dzakə**	**Ye**	**Nwabang**
Guinea fowl	[só]	[só]	[ʃó]
Water pot	[d͡zɔ̏]	[zɔ̏]	[t͡ʃɔ̏]
a specie of grass	[kə́pídàŋ]	[kə́bídàŋ]	[g͡bə́kə́dàŋ]
foot	[pù]	[bù]	[pù]
answer (v)	[kwə̀]	[gwə̀]	[kwə̀]
person	[ífɨ̀]	[ívɨ̀]	[ífɨ̀]
dust	[fò]	[vò]	[fò]
food server	[t͡ʃɛ̏]	[t͡sɛ̏]	[t͡ʃɛ̏]
open (v)	[bě]	[ɓě]	[bě]
give (v)	[tə̀]	[də̀]	[tə̀]

Other instances of dialectal difference I observed are drawn from earlier recordings of Dza speakers from the Ye region made by the Gospel Recording Network (GRN) in (1) and a recording of folk songs of a speaker from the Dzakə region which I made in 2018 in (2) below:

(1) a. *GRN Recording*

<min tsũpɨ be mɨnggba agã>

[mìn t͡sű̃pɨ̀ bê mɨ́ŋg͡bá àgǎ̃]

/mì-n   tű̃-pɨ̀   bê   mɨ́ŋ-g͡bá   à-gǎ̃/

1.SG-NPST   send-LOC   with   water-break   SPEC-big

'I will send a big flood'

b. *Present form of language*

<min tũpɨ be mɨnggba akã>

[mìn t͡ű̃pɨ̀ bê mɨ́ŋg͡bá àkǎ̃]

/mì-n   tű̃-pɨ̀   bê   mɨ́ŋ-g͡bá   à-kǎ̃/

1.SG-NPST   send-LOC   with   water-break   SPEC-big

'I will send a big flood'

(2) a. *Folk song*

<navo>

[nǎvò]

/nǎ   və̀   wə̄/

mother   GEN   3SG.POS

b. *Present form of the language*

<Na o>

[nà ộ̤]

/nǎ   wə̄/

mother 3SG.POS

Relative to the form of Dza in example (1) above, the current form of Dza shows lexical differences in terms of the devoicing of the velar stop [g] to a voiceless counterpart [k], and the fortition of the alveolar affricate [t͡s] to an alveolar stop [t]. In example (2), it shows a surface phonetic difference in terms of the mutation of the voiced labial – dental fricative [v] into a breathy vowel [o̤] which I speculate to be a post-debuccalized form of the voiced labial-dental fricative. This type of consonant change coincidentally supports the hypothesis that breathy phonation might be a vestigial trace of sound change in Dza.

3.3 Language Use and Multilingualism

My estimates are 70% population of Dza speakers are trilingual, speaking Dza, Hausa and one other language while 30% are bilinguals, speaking Dza and one other language mostly Hausa. Hausa is the language of wider communication in Northern Nigeria. In the Dza area, Hausa is used in the market, churches, and mosques. Dza has its prominence in homes and other social gatherings local to the Dza people. It is the official language of two African Traditional Religions: the *Nɨngbwi* cult and *Mə-akã* religion. English is the language of formal education and white-collar jobs. English is also used in new generation Nigerian indigenous churches. Dza speakers speak neighboring languages such as Bacama, Kwa, Bandawa, Wurbo et cetera. Typically, they learn these languages through intermarriage, religion, and other sociocultural interactions. Wurbo is a language worthy of mentioned within that region. It is a jukunoid language and the language of liturgy for a religion practiced among the Dza people and their neighbors. The religion is called *Mə Wurbo* alternatively *Mə Mbaya* in Dza. In this religion, new converts are obliged to learn the Wurbo language for communication with other members of the religion from neighboring communities and for ritual purposes.

3.4 Aspects of Dza phonetics and phonology

3.4.1 Consonant inventory

In Othaniel 2016, I presented that Dza has 31 consonant phonemes and that the consonant phonemes show symmetry in terms of voicing and at various places and manners of articulation. Table 3 below shows the inventory of consonant phonemes and some sounds in parenthesis.

Table 3:Inventory of Consonant Phonemes of Dza

	Labial		**alveolar**		**palatal**		**labio-velar**		**velar**		**glottal**
Implosive		(ɓ)		(ɗ)							
Plosive	p	b	T	d	c	ɟ	k͡p	g͡b	k	g	
Affricate			t͡s	d͡z	t͡ʃ	d͡ʒ					
Nasal	(m̥)	m		n		ɲ				ŋ	
Fricative	f	v	S	z	ʃ	ʒ					h
Approximant				l	(j̥)	j	ʍ	w			
					ɥ̊	ɥ					

The implosives [ɓ, ɗ] are common in the Bikwin–Jen languages except for Dza (see Norton and Othaniel 2020). In Othaniel 2016, I observed that the contrast between the voiced bilabial plosive [b] and the corresponding bilabial implosive [ɓ] is no longer phonemic in Dza. Similarly, there is no phonemic contrast between the voiced alveolar plosive [d] and the alveolar implosive [ɗ]. However, I presented that in the loss of phonemic contrast between the implosives [ɓ, ɗ] and their counterpart plosives [b, d], [b] substitutes [ɓ] in all instances but the reverse is not true. So also, [d] is a substitute of [ɗ] in all environments but [ɗ] is not a substitute of [d] in some instances. See (4) and (5) below for an illustration of the said restriction:

(4) a. [ɓút͡ʃì] ~ [bút͡ʃì] *arrow*
    b. *[ɓɨ̀] ~ [bɨ̀] *tsetse fly*

(5) a. [ɗɨ̀] ~ [dɨ̀] *take*
    b. *[ɗùɗú] ~ [dùdú] *tickle*

The voiceless bilabial nasal [m̥] occurs as a rare sound in Dza. Compared to other Bikwin – Jen languages, the voiceless bilabial nasal [m̥] occurs only in Dza (see Norton and Othaniel 2020). Even in Dza, it only occurs in the root [m̥mɨ́] for the number 'five' and other words related to that number. Dza also has a number of syllabic nasals that I interpreted as non-phonemic sounds of Dza (see Othaniel 2016).

The contrast between the voiced palatal approximants [j] and its voiceless counterpart [j̥] is weak and that they are in complementary distribution. The voiced palatal approximant [j] occurs before oral vowels but never nasal vowels, the voiceless approximant [j̥] on one hand occurs before nasalized vowels and never oral vowels. Table 4 below shows some example words:

Table 4: Complementary Distribution between /j/ and [j̥]

Oral Vowel		Nasalized Vowel	
[jə́]	*sp. of fish*	[j̥ə̃́]	*animal*
[jɑ̀]	*scatter (v)*	[j̥ɑ̃̀]	*to cross over*
[jí]	*rise (v)*	[j̥ĩ́]	*to soak*
*[jɨ̀ŋ]	*haze*	[j̥ɨ́ŋ]	*fish*
*[jə̀ŋ]	*gorila*	[j̥ə́ŋ]	*scorpion*
*[jə̀ŋ]	*leaf*	[j̥ə̀ŋ]	*tear (v)*

The lexical items marked with asterisks (*) in table 3.8 above are possible exception to the complementary distribution claimed here. Given the [C_N] environment these vowels occurred in the example words, it could be debated that these are nasalized vowels.

3.4.2 Vowels

As I pointed out in the introductory section of this study, Dza has nine (9) modal vowels in its inventory: /i, e, ɛ, ɨ, ə, a, u, o, ɔ/. The modal vowels /i, ɛ, ɨ, ə, a, u, ɔ/ can be articulated as either oral or nasal vowels and that the contrast is phonemic. Furthermore, the Dza modal vowels /i, e, ə, a, u, o/ are contrastive with a set of vowels which I impressionistically analyzed as breathy voiced vowels /i̤, e̤, ə̤, a̤, ṳ, o̤/ (Othaniel 2016). Table 5.1 below shows the oral vowels, Table 5.2 the nasal vowels, Table 5.3 shows the breathy vowels, while some comments on the distributional properties of Dza vowels are given after the tables below:

Table 5: **Dza vowel inventories**

Table 5.1: Oral Vowels

	front	*Central*	*back*
close	i	ɨ	u
	e		o
mid		ə	
	ɛ		ɔ
open		ɑ	

Table 5.2: Nasalized Vowels

	front	*central*	*back*
close	ĩ	ɨ̃	ũ
mid		ə̃	
	ɛ̃		ɔ̃
open		ɑ̃	

Table 5.3: Breathy Vowels

	front	*central*	back
Close	i̤	ɨ̤	ṳ
	e̤		o̤
Mid		ə̤	
Open		ɑ̤	

Distributional properties of vowels in Dza: Vowels do not occur word initially except for the high front unrounded vowel /i/ which occurs in words such as /íbě/ "suffering" and /ìd͡zwá/ "dog". Vowel lenght contrast is not phonemic in Dza and so far proves to be so in other Bikwin – Jen languages (see Harley 2020; Norton and Othaniel 2020; Othaniel 2017). Dza has an ATR vowel harmony between the +ATR vowels [e, o] and the -ATR vowels [ɛ, ɔ], where the two sets of vowels do not co-occur in the same words. A patent restriction related to the +ATR vowels [e, o] and the -ATR vowels [ɛ, ɔ] occurs in the nasal versus breathy voice quality features in Dza.

Table 6: Minimal pairs between oral, nasalized, and breathy vowels in Dza

Oral		**Nasalized**		**Breathy**	
/fì/	*Ring*	/fĩ̀/	*plant spike*	/fì̤/	*maternal uncle*
/pè/	*Surpass*	-	-	/pè̤/	*to traverse*
/hɛ́/	*to swell up*	/hɛ̃́/	*all*	-	-
/fɨ̄/	*Duck*	/fɨ̃̄/	*desire*	/fɨ̤̀/	*to pull*
/ʍə̀/	*Clay*	/ʍə̃̀/	*to rift off*	/ʍə̤̀/	*hunger*
/kà/	*to embrace*	/kã̀/	*to thank*	/kà̤/	*to mute*
/kù/	*to dust off*	/kũ̀/	*to shorten*	/kṳ̀/	*to belch*
/sò/	*mouth cavity*	-	-	/sò̤/	*elephant*
/tɔ̄/	*odour*	/tɔ̃̄/	*ostrich*	-	-

Table 6 above presents minimal pairs and triplets showing contrast between the three sets of vowels; oral, nasalized and breathy. The high mid +ATR vowels /e, o/ do not occur nasalized in Dza, but they do occur as breathy. On the other hand, the -ATR vowels [ɛ, ɔ]

3.4.3 Tone

In Othaniel 2016, I illustrated that Dza is a tonal language, and that tone is utilized for both lexical and grammatical contrast. I mentioned that Dza has three contrastive level tones which I tagged H, M and L. In addition, there are three contour tones: HL, LH and LHL. Below are examples of the lexical tones in Dza adapted from Othaniel 2020.

Table 7: **Lexical tonal contrast**

Tone Level	**Example word**		**Example word**		**Example word**	
H	/tá/	*to cry*	/kú/	*head*	/ɥ̥é/	*red monkey*
M	/tā/	*to press*	/kū/	*shade*	/ɥ̥ē/	*tamarind*
L	/tà/	*tobacco*	/kù/	*shake of*	/ɥ̥è/	*snail*
LH	/tǎ/	*father*	/kǔ/	belch	/ɥ̥ě/	*bed*

In Dza, tone is one of the morphological strategies for marking the contrast between the interrogative and the declarative mood, object pronouns and possessive pronouns, perfective and imperfective (habitual) aspect. In examples (3), the difference between the plain statement in ‘i’ and the question in ‘ii’ is the high-fall tone on the reduplicated verb root. While in (3) iii, the possessive pronoun in the sentence final position is marked by a mid-tone while the object pronoun in the middle of the sentence by a low tone. In the negated constructions in (3) iv and v, habituality is marked by the complex low-falling contour tone on the negation clitic /lò/. Example sentences with grammatical functions of tone are given in (3) below:

(3) i. /mə̀-n tə́-tə́/
2.SG-NPST go-go
'You will go'

ii. /mə̀-n tə́-tə̂/
2.SG-NPST go-go.INTER
'Will you go?'

iii. /mə̀-n t͡ʃī mì bê hò mī/
2.SG-NPST find 1.SG with bag 1.SG.POS
'You will find me together with my bag'

iv. /ò ŋwá lò/
2.SG.PFV drink NEG
'He did not drink'

v. /ò-ṇ̀ ŋwá lȭ/
2.SG.PFV-NPST drink NEG.HAB
'He does not drinks'

A fourth grammatical function of tone is marking locative nominal forms. See examples (3) vi below for examples:

vi. / té/ mountain /tê/ *on the mountain*
vii. /ŋwī/ *stream* /ŋwȉ/ *in the stream*

The morphological difference between the noun 'mountain' and the locative phrase 'on the mountain' in 'vi' above is given by the high-low contour tone. Similarly, in 'vii', the morphological difference between the noun 'stream' and the locative phrase 'in the stream' is given by the mid-low falling tone.

3.4.4 Syllable types and word structures in Dza

There are six phonological syllable types in Dza. Table X below shows the syllable types and example words. Syllables can have onsets and coda. Similarly, syllables can lack onsets, coda, or both. Dza prefers simple onsets against complex onsets. The complex onset type allowed in Dza is strictly a consonant followed by an approximant. It is either a plosive followed by a liquid as can be seen in Table 3 below or, a [Cw] or [Cj] structure. The syllable types consist of a tone bearing unit which is either a vowel or a syllabic nasal. In terms of the distribution of consonants within the syllable, all the consonants in Table x above do occur as onsets. Phonetically, the consonants /ŋ/, /n/, /w/ and /p/ can occur in the coda position. It should be noted that /-w/ is an enclitic in future negation construction as mentioned above. The only occurance of /p/ in the coda position is in a traditional chieftaincy title /záp.té/. The title is an alternative to /ívə̀/ which mean 'king'. This nickname, I propose is in accord with unusual phonology of interjections and ideophones in Dza. Meanwhile, in content words, /n/ usually occur in bisyllabic words as coda of the first syllable. And example word is /wāndē/ 'donkey'. This I propose is due to nasal assimilation.

Table 8: **Syllable types**

Syllable Patern	**Example word**	**Gloss**
V	/ɛ̀/	*broom*
	/ò/	*to fall*
CV	/bá/	*goat*
	/mə̀/	*to repair*
CVC	/pīŋ/	*egg*
	/sə̀ŋ/	*to filter*
CCV	/klɔ́/	*chief*
	/blō/	*watery*
Ṇ	/ǹ̩/	*1sg Subject pronoun*
	-	-
VC	/áŋ.kə̀lə̀/	*warning*
	ow ˧	*will not fall*

Typically, words are monosyllabic in Dza. Verb forms, apart from phrasal verbs, are either [V], [CV] or [CVC]. However, in the case of nouns, the above syllable types can combine to form two syllables, three syllables and four syllable words. Table x below shows the various possible word structures so far seen in Dza with examples. From the list of the word structure in Table 4, the syllable types [CV] and [CVC] occur without distributional restriction. They can occur in word initial, word medial and word final positions. However, the syllable types [V], [VC] and [Ṇ] are restricted to the word initial position. In fact, this restriction triggers adjustment whenever there are some morphological operations that could yield phonological structures different from the ones illustrated in tables x above and table x below. Gussenhoven and Jacobs (2017) presented two ways phonological adjustments could occur in a language: When words are borrowed from another language or, within the native vocabulary, when morphemes are combined, and they result to an ill-formed structure. For example, Dza does not allow diphthongs or vowel sequence as illustrated in the tables above. A native speaker of Dza pronounces the hausa word [a^{i}ki] 'work' as [ɛki]. This is an example of phonological adjustments that appear in loanwords.

Table x: Polysyllabic Word Structures In Dza

S/No	Structure	Example	Gloss
1.	**V.CV**	[íbě]	'suffering'
2.	**V.CVC**	[īlə̄ŋ]	'calabash'
3.	**CV.CV**	[bút͡ʃī]	'mud'
4.	**CV.CVC**	[sə̀kə́ŋ]	'k.o fish'
5.	**CVC.CV**	[t͡ʃə̀ŋkə́]	'k.o bat'
6.	**CVC.CVC**	[kīŋkə̄ŋ]	'husk'
7.	**Ṇ.CV**	[m̩̀pɛ̄̃]	'k.o grass'
8.	**Ṇ.CVC**	[ǹ̩dɨ̀ŋ]	'shank'

9.	**CV.CVC.CV**	[nàbɨ́ŋjù]	‘k.o grass’
10.	**CV.CV.CV**	[kúdə̀kó]	‘Sweet potatoes’
11.	**CV.CCVC.CCVC**	[t͡sɛ̀kwə̀ŋkwə́ŋ]	‘k.o falcon’
12.	**CVC.CV.CV**	[t͡ʃɨ̀nt͡ʃàlì	‘gold’
13.	**CV.CCV.CCV**	[gùlwêtwê]	‘k.o bird’
14.	**CV.CV.CVC**	[tə̀búd͡ʒɨ̀ŋ]	‘mist’
15.	**N̩.CV.CV**	[m̩̀bə̀gȕ]	‘dragonfly’
16.	**V.CV.CV**	[ìtɛ̀gɔ̀]	‘k.o bird’
17.	**CV.CVC.CV.CV**	[pùkə̀ntɛ̏nɛ̂]	‘k.o plant’
18.	**CV.CV.CV.CV**	[pàpànɔ̀nɔ̂]	‘k.o ant’
19.	**CVC.CV.CV.CV**	[t͡ʃə̆ŋkùlōlō]	‘centipede’

4 Hypotheses and the predictions

Firstly, in section 3.4 above I presented three sets of vowels viz., oral, nasal, and breathy vowels. My proposition here is that speakers of Dza do contrast vowel segments in terms of laryngeal settings at two points, the first is the neutral mode of phonation in which modal vowels are produced and the second is a non-modal phonatory setting in which what I called breathy vowels are produced. Modal phonation is associated with acoustic properties such as regular periodicity void of controllable noise, substantial overall acoustic intensity, medium range of fundamental frequency, intermediate spectral tilt values relative to non – modal phonation types across the phonation type continuum (see Gordon and Ladefoged 2001; Laver 2009). Whereas breathy phonation is associated with increased spectral noise at higher frequencies, decrease in overall acoustic intensity, lowered fundamental frequency and higher spectral tilt relative to modal phonation (Gordon and Ladefoged 2001; Garellek 2019). I anticipate that the modal and the breathy sets of vowels in Dza will contrast acoustically by some the regular sets of parameters for measuring voice quality. However, studies have shown that languages differ in the precise set of acoustic parameters they use in distinguishing phonation types (see Blankenship 2002; Gordon and Ladeforged 2001). It is therefore difficult to anticipate the exact cues Dza speakers use in making phonation contrast in vowels. However, if I measure several acoustic parameters used in measuring voice quality in several other languages, I expect some sets of acoustic parameters to capture the contrast between the two sets of vowels better than others.

Secondly, I presented in Othaniel 2016 that the modal-breathy phonation contrast in Dza is not phonemic because it disappears in speech. This phonological statement is based on my assumption that the modal – breathy contrast is not systematic phonetically. However, studies have shown that non-modal vowels are perceptually less robust than modal vowels due to the reduced salience of non-modal voicing (Gordon 1998, Blankenship 2002). Thus, my submission may be due to perceptual limitations. However, if my initial surmisal is true, without an explicit instruction given to the participants on the type of contrast observed in the production study, I expect variation in the outcome. By this I mean that speaker A may make the modal-breathy contrast between a set of minimal pairs, while speaker B may not, and I expect that this is not the result of gender differences as reported in Klatt and Klatt 1990. And I expect that the variation is not due to age or regional differences.

5 Methods

5.1 Speakers

This research is based on a production study which took place in Jimeta town, Adamawa state Nigeria. The participants are five males and five females Dza speakers with no known vocal disorders. However, two of the speakers were not audible during the production exercise, so their recordings were kept aside. The speakers included in this research are of the age range 25 – 47 years. All the speakers are residence in the Dza area, and they use the language in their daily lives. The participants were drawn from the three regions of Dza, viz., Ye, Dzakə and Nwabang areas. They all have some education; two of them are university graduates with bachelor's degrees, four are university students pursuing bachelor's degrees, while the other four are ND/NCE holders. None of them is a linguist. The speakers were transported from the Dza town to Jimeta for this production exercise. There was no accident recorded during their travels and during the production exercise. Table 9 below shows the speakers, their age and gender.

Table 9: Speakers, gender, age and region

Speaker	***Gender***	***Age***	***Region***
S1	F	46	Dzakə
S2	M	35	Dzakə
S3	F	25	Nwabang
S4	M	46	Ye
S5	M	25	Ye
S6	F	47	Ye
S7	M	30	Nwabang
S8	M	47	Dzakə

5.2 Wordlist

I prepared a questionnaire based on my previous auditory transcriptions of Dza. The questionnaire contains 33 lexical items of which eight (8) items are fillers. The questionnaire includes minimal pairs of: (a) voiced vs. voiceless onset consonants, and (b) modal and breathy vowels. Thirteen (13) of the words have voiceless onsets and breathy voicing on vowels. Seven (7) items are voiced onsets followed by modal vowels, and thirteen (13) are voiceless onsets followed by modal vowels. The thirty-three lexical items where organized based on semantic fields. For this study, I targeted four vowels /e, ə, a, o/ which I expected to contrast with /e̤, ə̤, a̤, o̤/. Before I had the previlege of learning how to use VoiceSauce, my initial intention was to measure only the voice parameters H1 – H2, H1 – A3 in Praat either manually or using Praat scripts. I was advised of the potential 'boosting' effect of F1 on the amplitude of H1 and H2 due to the proximity of F1 and H1 and H2 in high vowels (Ian Maddieson, per. comm). Thus, I decided to exclude high vowels in the current study.

5.3 Process

Each questionnaire item was recorded with three repetitions in isolation and then three times in the example sentence in (2) below. Explicit instructions were given to the participant to pronounce the example sentence all at once as a continuous speech where the speakers try to pay attention to their pronunciation of the target word in the frame sentence as if they were pronouncing it in isolation,

(2) <A fəng _______pɨnɨ>

/à	fə̀ŋ	______	pɨ̀nɨ́/
2.SG.IMP	scrape	target word	here

The recording was made at the sample rate of 48 kHz, 24 bits with Sound Devices 702 and a Shure TwinPlex TH53 omnidirectional head-worn condenser microphone complemented with FetHead Phantom preamplifier. Each speaker's long sound file was segmented according to individual questionnaire items. By that I mean each segmented sound file has the three iterations of the questionnaire item. And then I created TextGrid with three tiers for each of the segmented sound files in Praat (Boersma and Weenink 2011). In the first tier of the TextGrid I tagged the consonants and the vowels, in the second tier I tagged the words and the number of repetitions, and the third tier contains the English gloss of the word.

For consistency in marking vowel onset and offset, I followed the systematic convention proposed by Turk et al (2006). I relied primarily on visual spectral cues from zoomed out spectrograms to determine the location of boundaries. Afterwards, I zoom in on the waveform at the boundaries I marked and adjust where necessary. All questionnaire items are monosyllabic CV words. For a vowel preceded by the stops, viz., [b, d, p, t]; the release burst, onset of F2 and the transition of harmonic energy are the criteria I used in marking the onset boundary of the vowel, while F2 offset was what I used to mark the offset of the vowel. In words that have fricatives as onsets viz., [z, s, t], the onset of F2 at the offset of the frication energy is what I used to mark the onset boundary of the vowel segment. The words that have the approximants [ɥ, ɥ̥] as onset, I marked the boundary at the midpoint of transition from the approximant to the vowel.

After labelling the files in Praat, I used a Praat script written by Professor Jalaal Al-Tamimi to cut the individual iterations as separate sound files and their corresponding part of the TextGrid into an independent TextGrid file. The script also performed accurate F0 estimation for each speaker based on the two-passed method. The script was later modified to find special characters / "i̤, ṳ, e̤, o̤, ə, ə̤, a̤, ɥ, ɥ̥" which I used in the initial labelling to replace them with "ih, uh, eh, oh, ee, eeh, yw, hyw" respectively. The special characters failed to work with the program I chose to use for the voice quality analysis.

The sound files of the iterations and their corresponding TextGrid files were analysed in VoiceSauce (Shue et al 2011). In the parameter analysis settings in VoiceSauce, I chose Praat. For each speaker, I changed the minimum and maximum F0 values according to the accurate F0 estimation values generated for that speaker from the Praat script mentioned earlier. I allowed VoiceSauce to compute all the voice parameter measures in its default setting. By default, the program takes a measure of the various voice parameters at the interval of 1 msec across the duration of each sound file. VoiceSauce then generates the results according to the tagged labels in the corresponding TextGrid of the sound file. In this research, I am interested in studying the effect of breathy phonation over the vowel duration as one of the indications whether phonation type contrast is systemic in Dza or not. It will take me a lot of time to analyse the measurements at an interval of 1 msec across the duration of each vowel segment, considering I do that semi-manually. Thus, I tasked VoiceSauce to generate the results for each tagged segment in the TextGrid over nine (9) sub-segments. For each of the nine (9) sub-segments, VoiceSauce computes the mean scores of all the intervals in that sub-segment and generates the result as a text file. I then organized the results for all the speakers in one excel file and I analysed them by generating pivot tables.

5.4 Parameters to measure

The choice of acoustic parameters measured in this research is influenced by the psychoacoustic model of voice proposed by Kreiman et al. (2014) as presented in Garellek (2019). The psychoacoustic model of voice has four components; the first is the harmonic structure of the voice source spectrum (the sound produced by the vocal folds, before being altered by the vocal tract) measured by the parameters; *H1 – *H2, *H2 - *H4, *H4 - *H2kHz, *H2kHz - *H5kHz. The second is the inharmonic component of the source spectrum (i.e. noise) measured by HNR. The third is a set of two temporal components of the voice source (f0 and amplitude) and the transfer function of the vocal tract (Garellek 2019). Garellek noted that only the first three components are relevant for phonation type distinction. Thus, this research only focused on parameters of the first three components of the psychoacoustic voice model. The list of the acoustic parameters measured in this study are presented in Table 10.

Table 10:**List of acoustic parameters**

	Acoustic parameter	***Expected relationship***
Spectral tilt	H1-H2	Higher for breathy vowels than for the modal vowels
	H2-H4	Higher for breathy vowels than for the modal vowels
	H4 – H2 kHz	Higher for breathy vowels than for the modal vowels
	H2 kHz – H5 kHz	Higher for breathy vowels than for the modal vowels
	H1 – A3	Higher for breathy vowels than for the modal vowels
Spectral noise Noise	HNR35	Lower for breathy vowels than for the modal vowels
	CPP	Lower for breathy vowels than for the modal vowels
Fundamental frequency (F_0)	F_0	Lower for breathy vowels than for the modal vowels

Figure 3: List of acoustic parameters

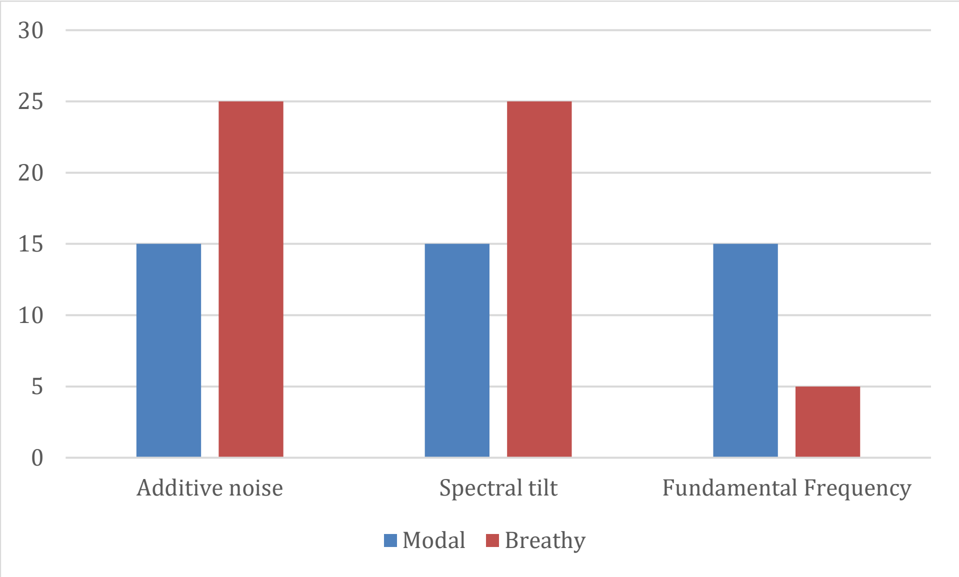


Figure 4: Summary of expected outcome

In the list of spectral tilt parameters in Table 3, the difference in amplitude between the first and second harmonics (H1 – H2) is the most widely discussed measure of spectral tilt in voice quality literatures (Klatt and Klatt 1990; Blankenship 2002; Esposito and Khan 2012; Wayland and Jongman 2003 etc). Physiologically, H1 – H2 correlates with the open quotient (OQ) of the glottal circle such that the larger the open quotient, the greater the amplitude of the first harmonic over that of the second harmonic (Holmberg et al 1995). For comparison purposes, I included H1 – A3, the difference in amplitude between the first harmonic and the prominent harmonic around the F3 frequency in the list of spectral tilt parameters for the purpose of comparison. H1 – A3 is utilized in various voice studies to describe the difference in spectral balance between phonation types (see Wayland and Jongman 2003; Keating et al 2021 for example). Figure 5 shows a sample harmonic spectrum for the four lowest-frequency harmonics (H1, H2, H3, H4), the harmonic nearest 2000 Hz, and harmonics nearest the formant frequencies F1, F2, and F3 (A1, A2, A3) from Keating et al. (2021). The harmonic nearest to 5000Hz is not indicated in the figure since spectrum only covers 0 – 3000Hz.

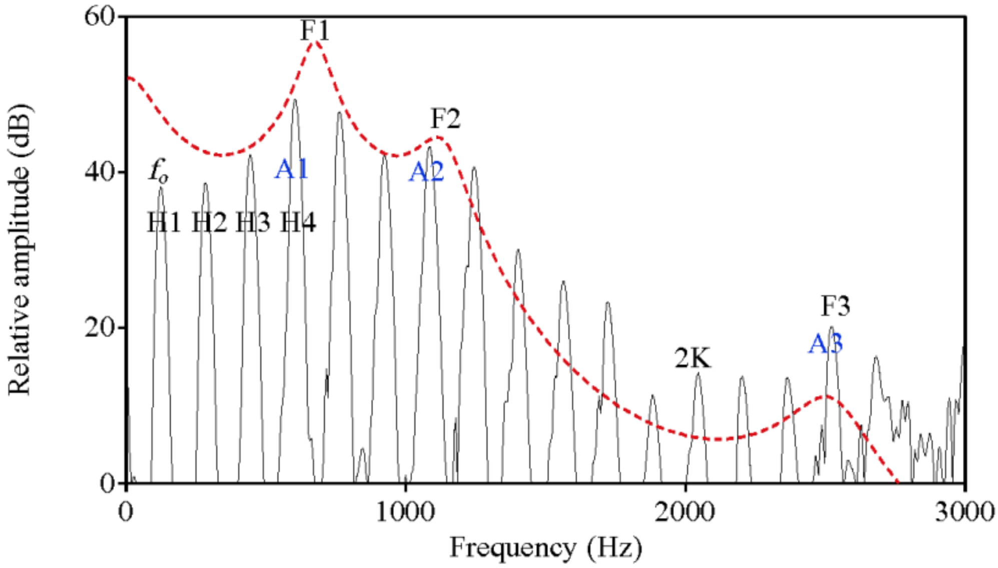


Figure 5: Sample harmonic spectrum showing spectral tilt parameters and F0

For the spectral noise measurement, I measured two parameters HNR35 and CPP. By default, voicesauce measures HNR at different frequency ranges, in this research I only considered the HNR measurement of the frequency range 0 – 3500Hz. In contrast, the CPP measurement covers the entire frequency range (Shue et al. 2011).

6 Results

The results are presented with simple bar charts, longer bars indicate higher values. The results are organized in three main sections: spectral tilt (*H1 – *H2, *H2 - *H4, *H4 - *H2kHz, *H2kHz - *H5kHz, *H1 - *A3), spectral noise (HNR35, CPP) and the F0. Each speaker's result presented here are averages of mean differences at the first portions, fifth portions and the ninth portions of the various segments computed by VoiceSauce as explained in section 5.3 above. Since the first portion contains the average mean difference of the onset of the vowel segment in every case, I shall refer to it as the onset portion. For similar reason, I shall refer to the fifth and the ninth portions as the midpoint and offset portions respectively. The asterisk (*) denotes that the spectral magnitudes (H1, H2, H4, H5 and A3) are corrected in voicesauce for the effect of the frequencies and bandwidths of adjacent formants (see Shue et al 2011). The measures for spectral tilt parameters: *H1 – *H2, *H2 - *H4, *H4 - *H2kHz, *H2kHz - *H5kHz, *H1 - *A3 are presented with a numeric scale from – 5 to 30 dB. Spectral noise parameter measures: CPP and the HNR35 are presented on a numeric scale from 0 – 60 in dB. The F0 track is presented with a numeric scale from 0 – 250 in Hz.

6.1 Spectral tilt

6.1.1 *H1 – *H2

It is expected that the *H1 – *H2 value should be higher in the breathy sets than in modal phonation. Figure 4 below shows a summary of all speakers' measurements of *H1 – *H2.

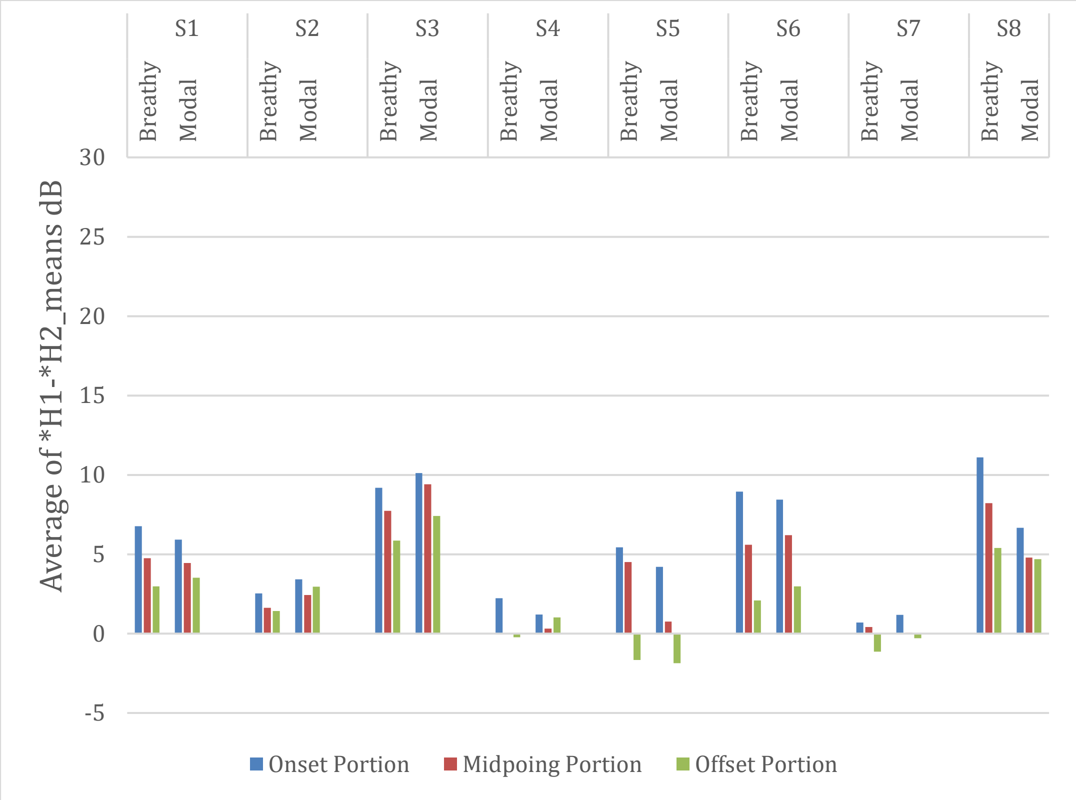


Figure 6: *H1-*H2 All speakers summary

The *H1 - *H2 measurement results show that the eight speakers contrast the two sets of vowels in different ways. In the onset portion of the vowels, the results for speakers 1, 4, 5, 6, and 8 indicate slightly higher H1 – H2 values in the breathy segments than their modal minimal pairs as expected. The results for speakers 2, 3 and 7 are rather in the opposite direction, with *H1 - *H2 values slightly higher in the modal segments than those of their breathy minimal pairs. As mentioned in section 2 above, when the H1 – H2 value is lower relative to the modal segments, it indicates creakiness (Garellek 2019). In the midpoint portion, the results for speakers 5, 7 and 8 show higher *H1-*H2 values in the breathy set than in the modal set. Meanwhile, the results for speaker 1 showed virtually a flat *H1 - *H2 values for the breathy and the modal segments, and the results for speakers 2, 3, 4 and 6 showed the modal segments have slightly higher values that their breathy counterparts again suggesting creakiness. At the offset portions, the results for all the speakers showed higher *H1 - *H2 values in the modal segments than in their breathy minimal pairs except for speakers 5 and 8 whose results showed slightly higher values in the breathy segments relative to the modal ones.

In summary, the results of the *H1 - *H2 measurement for speakers 5 and 8 indicate higher values for the breathy sets in all positions across the duration of the vowels as expected. On the contrary, the results for speakers 2 and 3 showed higher values for the modal sets relative to their breathy counterparts throughout the three portions measured across the duration of the vowels. Speakers 4 and 6 results at the onset portions showed higher values for the breathy sets than the modal sets, and higher values for the modal sets at the midpoint and offset portions. For speaker 3, the results showed higher values for breathy sets at the onset portion, flat values for both breathy and modal sets, and then higher values for the modal sets at the offset portions.

6.1.2 *H2 - *H4

The expected results for the *H2 - *H4 measurements is like *H1 - *H2 above. Higher *H2 - *H4 values is expected for the breathy sets relative to the modal sets. Figure 5 below gives the summary of *H2 - *H4 measurement results for all the speakers. The results for speakers 1, 2 3, 4, and 6 at the onset portions showed slightly higher values for breathy sets compared to their modal counterparts as expected. For speakers 7 and 8, the values are nearly flat at the onset portions of both the breathy and the modal sets. This is also expected, that some speakers will not contrast between the two sets of vowels. At the midpoint portion, the results for speakers 1, 3, 4, and 6 showed higher values for the breathy sets compared to their modal counterparts. The results for speakers 2 and 7 showed nearly flat values for the breathy and the modal segments. Meanwhile, the results for speakers 5 and 8 showed higher values for the modal sets relative to their breathy counterparts. At the offset portion, the results for speakers 1, 2, 3, 4 and 6 show higher values for the breathy as expected, while for speakers 5 and 7, the results show slightly higher values for the modal sets. The result for speaker 8 shows flat values for both sets.

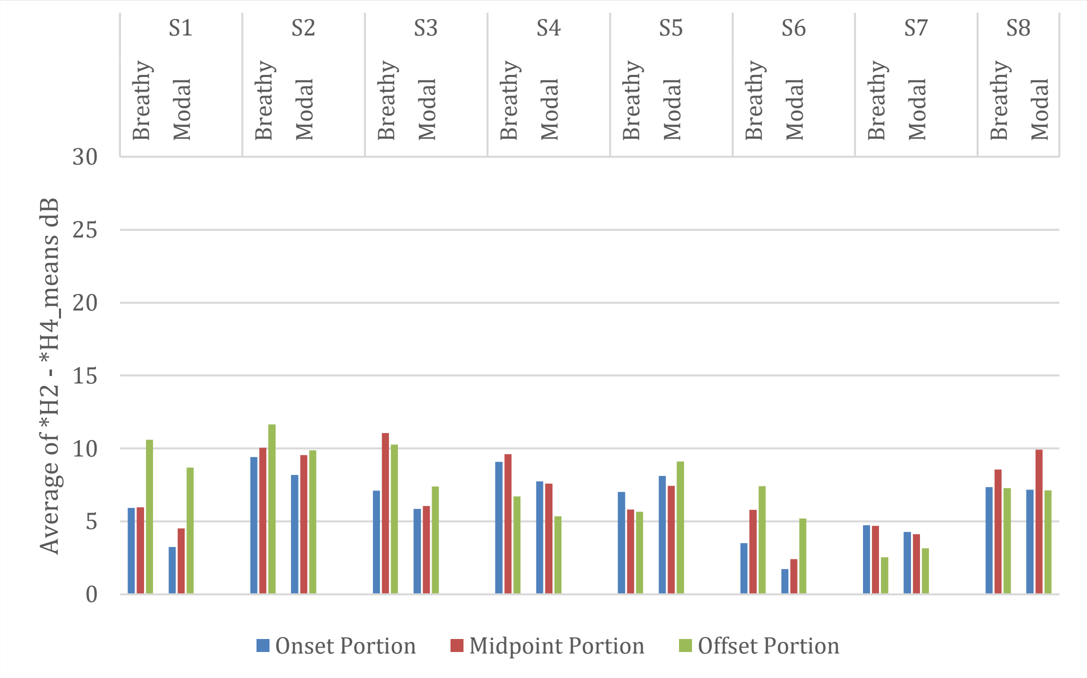


Figure 7: *H2-*H4 All Speakers Summary

6.1.3 *H4 - *H2kHz

The *H4 - *H2kHz measurement results in Figure 8 show no contrast between the two sets of vowels for speakers 1, 2 3, 4, 6 and 8 at the onset portion, while for speakers 5 and 7, the results showed slightly higher values for the breathy sets than the modal sets as expected. However, at the mipoint portions, the results showed higher values for the breathy sets than the modal ones for speakers 1, 2. 5, 6. 7, and 8. The result for speaker 3 shows higher values than the breathy set while the result for speaker 4 show a flat values for both the breathy and the modal sets.

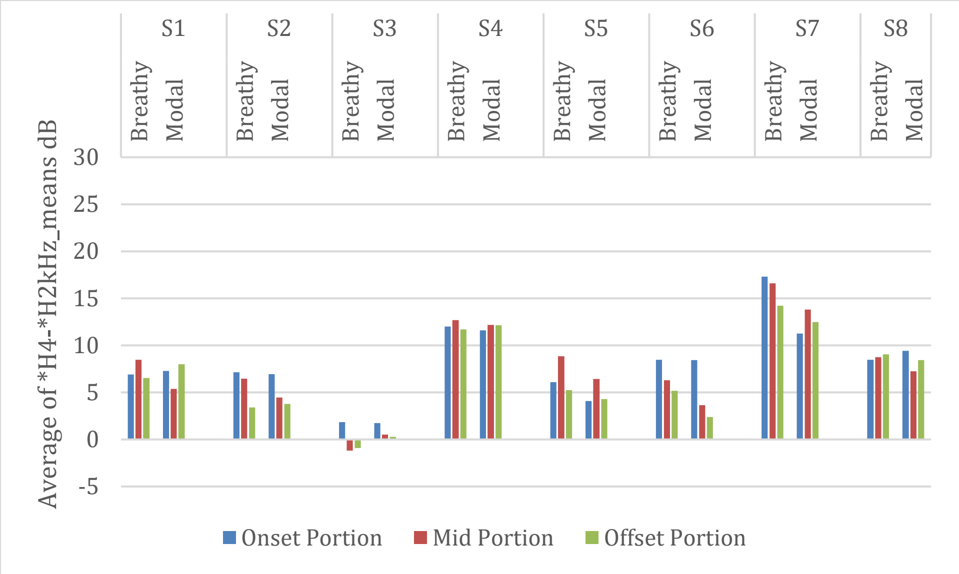


Figure 8: *H4-*H2kHz All Speakers Summary

At the offset portion, the results for speakers 1 and 3 show higher values for the modal sets than the breathy sets, while the values are flat for speakers 2, 4 and 8. The results for speakers 5, 6 and 7 show higher values for the breathy sets than their modal counterparts.

6.1.4 *H2kHz -*H5kHz

Figure 9 show the *H2kHz - *H5kHz measurement results. At the onset portion, the results for speakers 1, 5, and 8 no contrast between the two sets of vowels, while the results for speakers 2, 3, and 4 show higher values for the breathy sets than the modal sets of vowels.

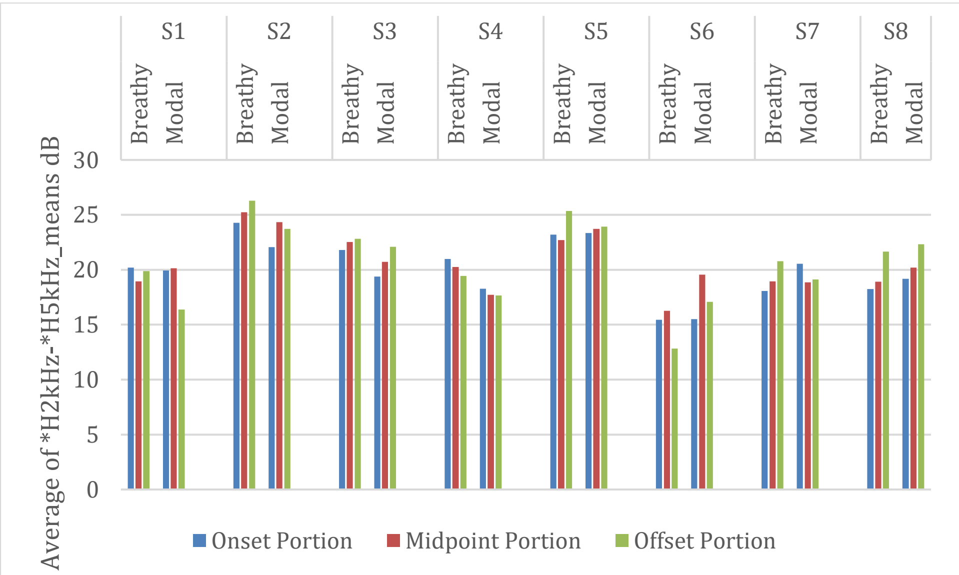


Figure 9: *H2kHz -*H5kHz All Speakers Summary

At the midportion, the results for speakers 2, 3 and 4 show slightly higher values for the breathy sets than the modal sets, no contrast for speaker 7 and slightly higher values for modal sets in the results of speakers 1, 5, and 8. At the offset, the result for speakers 1, 2, 4, 5, and 7 show slightly higher values for the breathy sets than for the modal sets. The results for speakers 3 and 8 show no contrast at the offset portions while speaker 6 result show a slightly higher value for the modal vowels than the breathy ones.

6.1.5 *H1-*A3

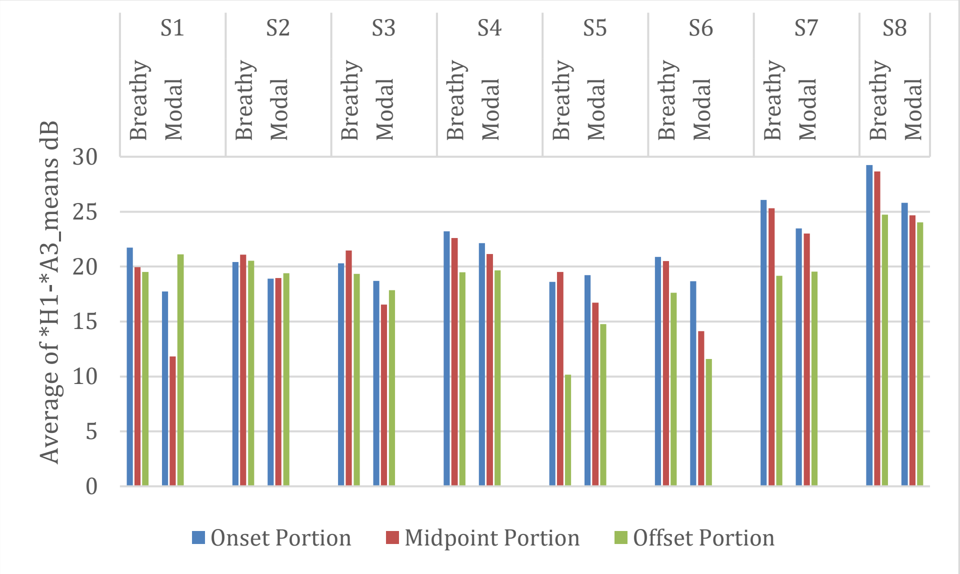


Figure 10: *H1-*A3 All Speakers Summary

The results for *H1 - *A3 measures at the onset portion show slightly higher values for the breathy sets in all speakers' results except for speaker 5 whose result show no contrast between the two sets of vowels. At the midportion, the results show that the *H1 - *A3 measures distinguished between the two sets of vowels for all speakers. The breathy sets have higher values than the modal vowels as expected. At the offset portions, the results for speakers 1 and 5 show higher values for the modal segment than the breathy, while the results for speakers 4, 7 and 8 show no contrast between the two sets of vowels. The results for speakers 2, 3, and 6 show slightly higher values for the breathy sets than for the modal ones.

6.2 Spectral Noise

Increased noise is expected at the higher frequencies in the breathy vowels than in the modal vowels. Thus, the harmonic to noise ratio for the breathy sets is expected to be less, and higher in the modal vowels. The result is rather in the opposite direction for all speakers. The HNR35 measurement values are higher for the modal set of vowels than the breathy vowels. Figure 11 below shows the summary results of the CPP measurement for all speakers while Figure 12 shows the results for the HNR35 measures.

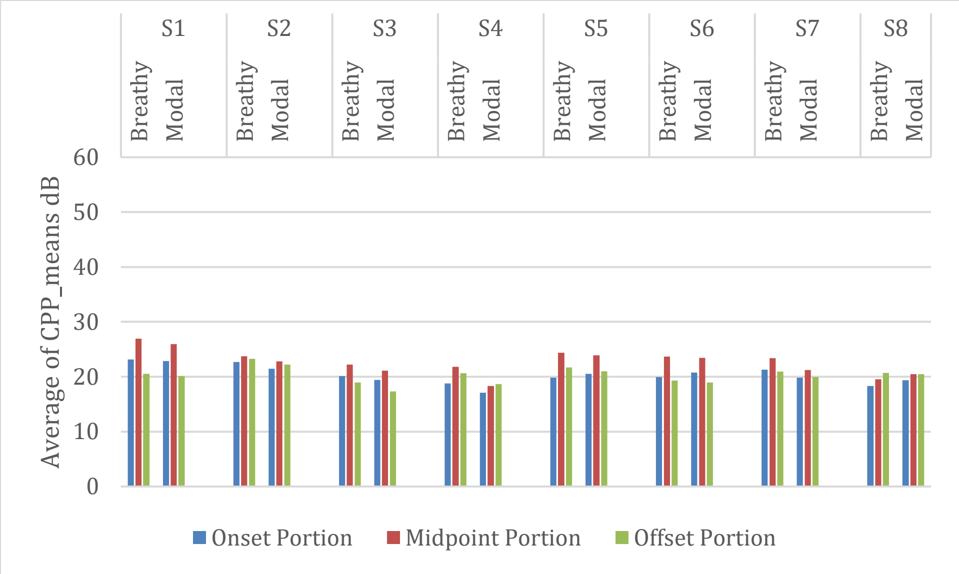


Figure 11: CPP All Speakers Summary

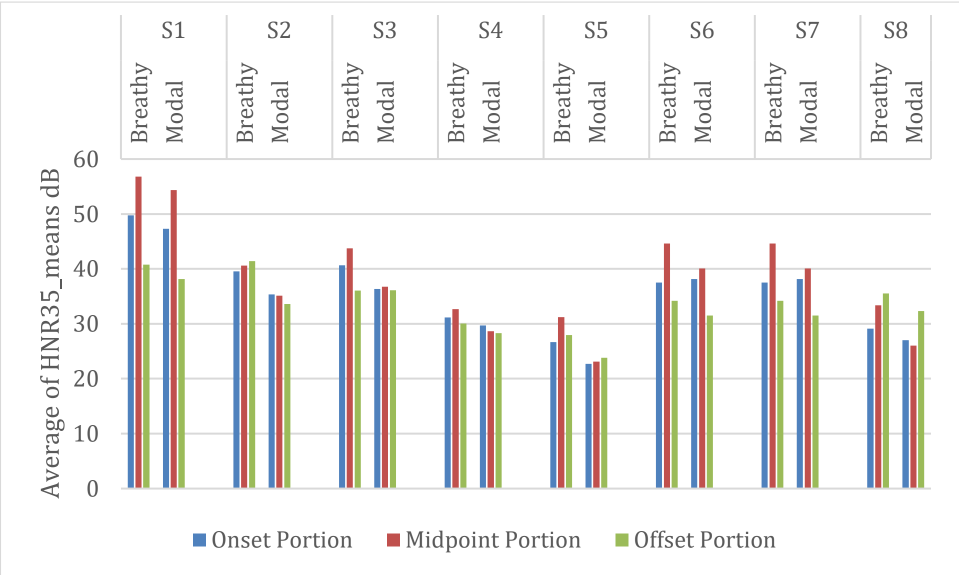


Figure 12: HNR 35 All Speakers Summary

6.3 F_0 track

The results for the f_0 track measurements are shown in Figure 13. As mentioned, breathy phonation is characterized by a lowered f0 relative to the modal phonation. Thus, the expectation is that the breathy sets will have lesser values compared to their modal counterparts. The f_0 measures results at the onset and midpoint portions of the vowels show higher values for the modal sets than the breathy sets as expected.

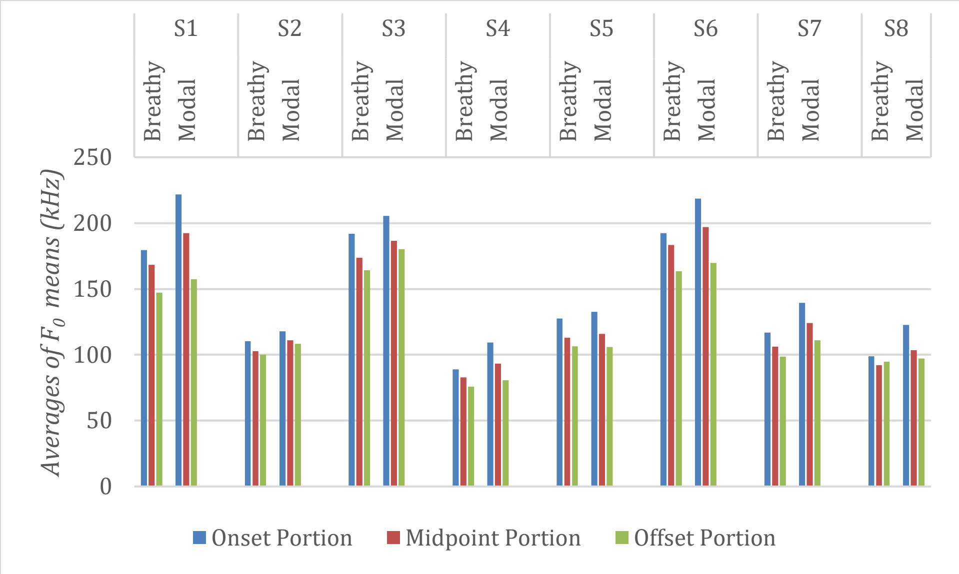


Figure 13: F_0 track

7 Discussion

7.1 Acoustic correlates of modal and breathy vowels in Dza

First, it is important to state which of the acoustic parameters distinguished between the two sets of vowels better according to the results. The nature of the contrast will be discussed afterward. The results for the spectral tilt measurements indicates that the lower frequency harmonic measures *H1 - *H2, and *H1 - *A3 distinguished between the two sets of vowels /e, ə, a, o/ and /e̤, ə̤, a̤, o̤/ better than *H2 - *H4 and *H4 - *H2kHz, *H2kHz - *H5kHz. For most speakers, the measures of prominent harmonics at higher frequencies (*H4 - *H2kHz, *H2kHz - *H5kHz) contrast the two sets of vowels with less magnitude when compared with *H1 - *H2, *H2 - *H4 and *H1 - *A3.

Furthermore, the *H2-*H4 measures indicate that the contrast between the two sets of vowels persists through the entire duration of the vowels for all speakers. However, when we compare with *H1 - *H2 and *H1 - *A3, the results showed the longer bars for most of the speakers are at the offset portions of the vowels and the difference between the two sets of vowels is much larger at the offset too. This means the difference in depth of spectral tilt is larger at the offset portion of the vowel than other parts of the vowel. However, sustaining the configuration of the vocal folds for non-modal phonation requires extra effort, normally the contrastive feature of non-modal phonation diminishes across the duration of the vowel lessens. This is already identified as one of the possible explanations why some languages localized non – modal phonation at the first portion of the vowel duration (See Gordon 1998, Blankenship 2002). Consequently, one expects larger difference at the onset portion rather than at the offset portion of the segments being compared.

From the results of the two measures of spectral noise parameters, the HNR35 differentiates between the two sets of vowels better than the CPP. As earlier noted, the CPP takes a measure of spectral noise across the entire frequency ranges while the HNR35 measures the harmonic to noise ratio at a specific frequency range of 0 – 3500 Hz.

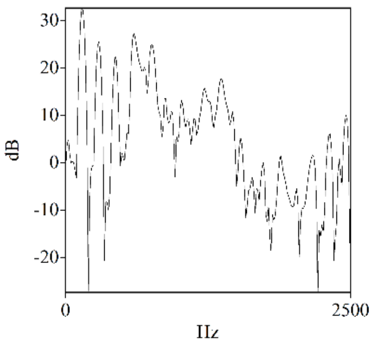
Lastly, the F0 measures also contrast between the two sets of vowels as shown in Figure 13.

7.2 Nature of modal versus non-modal phonation contrast among Dza speakers

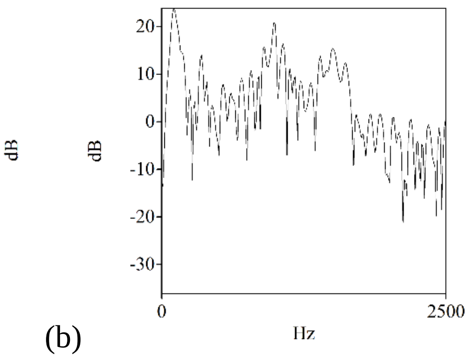
The results show a lot of variations among the speakers that participated in the study. The variations do not indicate a pattern in terms of region, sex, or age of speakers. The *H1-*H2 and *H1 - *A3 results for speaker 8 show higher difference in spectral dept in the breathy sets than the modal sets across the three portions of the vowels. For speaker 1, the difference in spectral tilt is higher for the breathy set at the first two portions of the vowel while at the offset portion the difference is higher for modal sets than the breathy. For the rest of the participants, the *H - *H2 and the *H1 - *A3 results are not similar. However, for most of the speakers the results for the two parameter measures at the onset portions indicate higher difference in spectral tilt for the breathy sets than for the modal sets. While the results of the measures of the two parameters *H - *H2 and *H1 - *A3 at the midpoint and offset portions vary for those speakers.

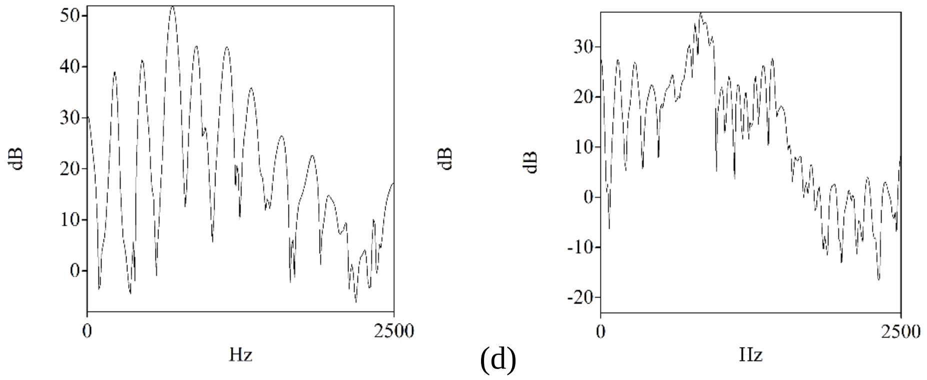
I predicted variation, but this is not exactly the pattern of variation I predicted. My prediction is variation based on the presence or absence of breathy phonation in vowels. I expected some speakers will distinguish between the two sets of vowels by phonation type while others will not. Instead, I observed at least three different patterns which the speakers distinguished between the two sets of vowels. I will present FFT spectrums which clearly show the first and the second harmonics to give an example with minimal pairs from different speakers to illustrate the different ways the speakers distinguished between the two sets of vowels. The spectral slices were taken 25ms into the vowel, at the midpoint region and at 25ms before the vowel offset. Where the pattern of contrast is similar at the onset region of the vowel, at the midpoint and the offset region, I will present spectral slices taken at the onset and the offset regions. Where the patterns differ at the three positions, I will present spectral slice from all the three positions.

(a)



(b)

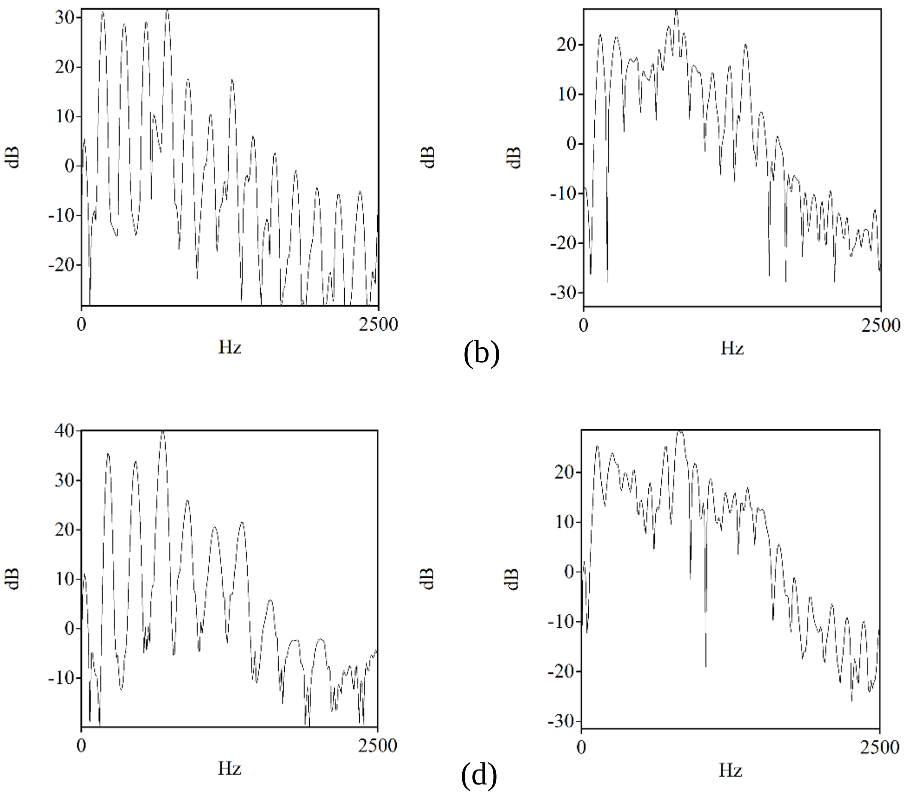




(c) (d)

Figure 14: **FFT spectrums of minimal pairs /pə/ and /pə̤/ produced by speaker 6 taken at 25ms into the vowels and 25ms before the offset**

In Figure 14, the first two spectrums (a) and (b) are taken from a breathy segment /pə̤/ "to pierce" produced by speaker 6. The two spectrums show the first harmonic having higher amplitude than the second harmonic. Spectrums (c) and (d) are taken from a modal segment /ṇpə/ produced by same speaker, the first two harmonics virtually have similar amplitudes. This is an example of my expectation of how modal and breathy vowels should be distinguished. In the case where the two sets of vowels do not contrast, I expected the case of speaker 1 in Figure 15.



(a) (b)

(c) (d)

Figure 15: **FFT spectrums of minimal pairs /pə/ and /pə̤/ produced by speaker 1 taken at 25ms into the vowels and 25ms before the offset**

In Figure 15, the first two spectrums (a) and (b) are taken from a breathy segment /pə̤/ "to pierce" produced by speaker 1. Spectrums (c) and (d) are taken from a modal segment /ṇpə/

produced by same speaker 1. When we compare between the two segments, there is not much difference between the first two harmonics. My predicted variation is the type of relationship between speaker 8 and speaker 1.

However, I observed three other patterns of variations. The first type of variation I observed is alternation of phonatory quality between the minimal pairs. This is the case where the supposed modal segments appear to be breathier than the breathy segments in some speakers' data, while in other speakers' data, the breathy segments are breathier. The second pattern of variation I observed is between the three iterations of individual items of the questionnaire. I observed cases where the three iterations have different phonatory qualities relative to the modal iterations. The fourth type of variation I observed is variation of phonatory quality across the duration of the vowel. This is the situation where the phonatory quality of the breathy segment, relative to the modal minimal pair, is different at the beginning of the vowel, different at the midpoint region and different at the offset region.

In summary, there was no systematic pattern observed in the way the speakers distinguished between the two sets of vowels. Rather, speakers even do different things between iterations of the same word.

7.3 Interpretation of phonation contrast

With reference to the continuum of phonation types in Figure 1 above by Gordon and Ladefoged (2001), modal phonation lies somewhere between the breathy and creaky phonation types. In terms of spectral tilt measurements as mentioned earlier, higher values are expected for segments with breathy phonation compared to segments with modal phonation. Similarly, higher values are expected for the modal phonation type compared to the creaky phonation type. Thus, when comparing between two sets of vowels, it is important to define the relationship between the two. For this reason, in several literatures of voice quality studies, the spectral tilt measures are often interpreted hand in hand with the spectral noise measures (see Garellek 2019; Blankenship 2002). In comparing two sets of vowels, the result is interpreted as breathy versus modal phonation contrast when one of the two sets show higher spectral tilt values, and less HNR values. However, if one of the two sets show higher spectral tilt values, and higher HNR values, the result is interpreted as modal versus unconstricted creaky phonation contrast. Similarly, if the results for one of the two vowel sets show less spectral tilt values, and less HNR values, the result is interpreted as modal versus constricted creaky phonation (Garellek 2019). Figure 14 is a summary taken from Garellek (2019) for illustration purposes:

	Vowel A	*Vowel B*	Vowel A′	*Vowel B′*
H1-H2	10	5	10	5
HNR	20	10	10	20
Interpretation	A has higher tilt, *less* noise than B. A = modal, B = creaky		A′ has higher tilt, *more* noise than B′. A′ = breathy, B′ = modal	

Figure 16:Sample H1–H2 and HNR values for two groups of Vowels A vs. B and their interpretations

Following this method of interpretation, the results of the *H1 - *H2 and *H1 - *A3 spectral tilt measurements, the HNR35 results in section 6 above suggest instead a contrast of modal versus creaky phonation for all speakers instead of a contrast of breathy versus modal phonation. In this way, what I have labelled as breathy vowels in Dza will be modal vowels instead, while what I have labelled as modal vowels will be creaky vowels instead. However, the results for the f0 tract presented in section 6.3 do not match this interpretation. Various publications on voice quality, including the psychoacoustic voice model, suggest a lower f0 for creaky phonation (Garellek 2019; Gordon and Ladefoged 2001). The results in section 6.3 showed the breathy sets of vowels have lowered f0 compared to the modal sets. Secondly, this interpretation may be valid when we restrict the definition of modal voice in articulatory terms. However, if we extend the definition to mean the default voicing in vowels, then the interpretation is problematic since what I called the modal vowels are the default vowels in Dza. The multiple patterns of variations observed in the study make the interpretation a little difficult. Nevertheless, a possible interpretation along the phonation type continuum is that what I called breathy phonation fall in between typical breathy vowels with higher noise with modal voice. This is often referred to as lax voice in voice quality studies (Kuang and Keating 2014; Garellek 2019).

8 Conclusion and further studies

The goal of this study is to investigate the acoustic correlates of modal versus breathy phonation in Dza and to determine to what extent is the modal versus breathy phonation contrast systematic among Dza speakers. The acoustic parameters *H1 - *H2, *H2 - *H4, *H4 - *H2kHz, *H2kHz - *H5kHz, *H1 - *A3, HNR35, CPP and F0 tracts were measured. The results of the analysis by VoiceSauce show that *H1 - *H2, *H1 - *A3, HNR35 and F0 distinguished between the two sets of vowels better than the other parameters. However, the contrast was not systematic. There was a lot of variations between the speakers, and even between repetitions of utterances by one speaker. The nature of the variation from this production study raised my curiosity as to how the speakers of Dza will distinguish between breathy, modal, and creaky vowels in a perception study. Another question from this study is whether the contrast between the two sets of vowels is not just about the phonation types but that it may also be a question of tonegenesis as earlier hypothesized. I believe a study of tone in Dza will contribute in explaining the variations observed in this study. Furthermore, I would like to improve my methods by using a more ecumenical statistical model in my analysis.

9 Works Cited

Blankenship, Barbara. 2002. "The timing of nonmodal phonation in vowels." *Journal of Phonetics* 163-191. doi:https://doi.org/10.1006/jpho.2001.0155.

Esling, John H, Scott Moisik, Allison Benner, and Lise Crevier-Buchman. 2019. *Voice Quality The Laryngeal Articulator Model* . Cambridge: Cambridge University Press.

Esposito, Christina M, and Sameer ud Dowla Khan. 20012. "Contrastive breathiness across consonants and vowels." *Journal of the International Phonetic Association* 123-143.

Garellek, Marc. 2019. "The phonetics of voice." In *The Routledge Handbook of Phonetics*, edited by William F Katz and Peter F Assmann, 75-106. Routledge.

Gordon, Matthew. 1998. "The Phonetics and Phonology of Non-modal Vowels: A Cross-Linguistic." *The Annual Proceedings of the Berkeley Linguistics Society.* eLanguage. 93-105. doi:https://doi.org/10.3765/bls.v24i1.1246.

Gordon, Matthew, and Peter Ladefoged. 2001. "Phonation types: a cross-linguistic overview." *Journal of phonetics 29(4)* 383–406.

Harley, Matthew. 2020. "Aspects of the Phonology and morphosyntax of Kyak, an Adamawa language of Nigeria." *Languages in Africa 1(3)* 373 - 404.

Holmberg, Eva B, Robert E Hillman, Joseph S Perkell, Peter C Guiod, and Susan L Goldman. 1995. "Comparisons Among Aerodynamic, Electroglottographic, and Acoustic Spectral Measures of Female Voice." *Journal of Speech and Hearing Research, Volume 38* 1212 - 1223.

Idiatov, Dmitry, and Mark Van de Velde. 2019. "Bena-Mboi is Benue-Congo." *Paper presented at the Adamawa Conference, Mainz, Germany.*

Keating, Patricia, Jianjing Kuang, Marc Garellek, Christina Esposito, and Sameer Khan. 2021. "A Cross-Language Acoustic Space for Vocalic Phonation Distinctions ." *Semantic Scholar.* Accessed April 7, 2022. https://www.semanticscholar.org/paper/A-CROSS-LANGUAGE-ACOUSTIC-SPACE-FOR-VOCALIC-Keating/9cffbf5a4437f243bf0973d897e28ff744ad8c29.

Klatt, Dennis H, and Laura C. Klatt. 1990. "Analysis, synthesis, and perception of voice quality variations among female and male talkers." *The Journal of the Acoustical Society of America 87, 820.* doi:https://doi.org/10.1121/1.398894.

Kleinewillinghöfer, Ulrich. 1996. "Relationship between Adamawa and Gur languages: The case of Waja and Tula." *Cahiers Voltaïques — Gur Papers 1.* 25–45.

Kreiman, Jody, and Bruce R. Gerrat. 2002. "Jitter, shimmer, and noise in pathological voice quality perception." *The Journal of the Acoustical Society of America* 57 - 61.

Kreiman, Jody, Bruce R Gerratt, Marc Garellek, and Robin, Zhang, Zhaoyan Samlan. 2014. "Toward a unified theory of voice production and perception." *Loquens.* doi:10.3989/loquens.2014.009.

Kuang, Jianjing, and Patricia Keating. 2014. "Vocal fold vibratory patterns in tense versus lax phonation contrast." *The Journal of the Acoustical Society of America 136, 2784.* doi:https://doi.org/10.1121/1.4896462.

Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages.* Oxford: Blackwell Publishers.

Laver, John. 1980. *The phonetic description of voice quality.* Cambridge: Cambridge University Press.

Maddieson, Ian. 2018. "Phonetics and African languages." In *The Languages and Linguistics of Africa,* edited by Tom Güldemann, 546-591. Berlin: Walter de Gruyter.

Norton, Russell, and Nlabephee Kefas Othaniel. 2020. "The Jen language cluster: A comparative analysis of wordlists." *Languages in Africa 1(3)* 17 - 99. doi:10.37892/2686-8946-2020-1-3-17-99.

Othaniel, Nlabephee Kefas. 2017. "A Phonological Comparative Study of the Jen Language Cluster." *TCNN B.A Thesis.* Jos, August.

—. 2016. "Jenjo Phonology Statement." Unpublished, 15 May.

Samarin, William J. 1971. "Adamawa - Eastern." In *Current Trends in Linguistics: Linguistics in Sub-Saharan Africa,* by Thomas A Sebeok, 213-244. Paris: Mouton.

Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2009. "VoiceSauce: A program for voice analysis." *The Journal of the Acoustical Society of America.* doi: https://doi.org/10.1121/1.3248865.

Suire, Alexandre, Arnaud Tognetti, Valérie Durand, Michel Raymond, and Melissa Barkat-Defradas. 2020. "Speech Acoustic Features: A Comparison of Gay Men, Heterosexual Men, and Heterosexual Women." *Archives of Sexual Behaviour* 2575–2583. doi:https://doi.org/10.1007/s10508-020-01665-3.

Turk, Alice, Satsuki Nakai, and Mariko Sugahara. 2006. "Acoustic Segment Durations in Prosodic Research: A Practical Guide." In *Methods in Emperical Prosody Research*, edited by Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Augurzky Petra, Ina Mleinek, Nicole Richter and Johannes SchlieBer, 1 - 28. Berlin: Walter de Gruyter.

Wayland, Ratree, and Allard Jongman. 2003. "Acoustic correlates of breathy and clear vowels: the case of Khmer." *Journal of Phonetics* 181–201.

10 Appendix I: Dza Breathy vs Modal Vowel Elicitation Wordlist

Frame:

<A fəng ________pɨnɨ>

/à fə̀ŋ _______ pɨ̀nɨ́/

2.SG.IMP scrape here

If the participant imagined someone wrote the target word on a desk; in that case the frame is saying scrape (target word) here. This should fit with all the items in the wordlist.

The participants should call each word in isolation three times and then use the word in the frame and repeat three times. It depends on the speaker, but my guess is, it should take them less than twenty (20) minutes each.

Wordlist:

The wordlist contains thirty-three (33) items, I tried to sort the items according to semantic domains. I added empty rows to be a sort of demarcation between the groups of semantically related items. There are seven (7) minimal triplet sets, six (6) minimal pairs in the wordlist. Table 1 is the tentative wordlist, Table 2 contains the minimal triplet sets, while Table 3 contains minimal pairs not contained in the minimal triplets in the wordlist.

Table 1: Wordlist

S/No	Word	Hausa	Dza
1.	‘to be tight’	‘yi matsi’	/tə̀/
2.	‘to pass’	‘wuce’	/pē/
3.	‘to hurry’	‘yi sauri’	/də̀/
4.	‘to see’	‘gani’	/bə̀/
5.	‘to give’	‘bayar’	/tə̤̀/
6.	‘to destroy’	‘rushe’	/dè/
7.	‘to pierce’	‘huda’	/pə̤̀/
8.	‘to spill’	‘zuba’	/tè/
9.	‘to protrude’	‘yi ɓulli’	/zò/
10.	‘to read’	‘karanta’	/sà̤/
11.	‘to answer’	‘amsa’	/kwə̤̀/
12.	‘to be lazy’	‘yi kiwuya’	/bē/
13.	‘to cover’	‘rufe’	/kì/
14.	‘to belch’	‘yi gyasa’	/kṳ̀/
15.	‘to seize’	‘kwace’	/kwə̀/
16.	‘to exchange’	‘yi musanya’	/pjə̀/
17.	‘gamba grass’	‘gamba’	/ǹ̩pə̀/
18.	‘k.o grass’	‘irin ciyawa’	/tḛ̀/

S/No	Word	Hausa	Dza
19.	‘to winnow’	‘yi bakace’	/pē̤/
20.	‘to sieve’	‘yi tankade’	/kù/
21.	‘soup’	‘miya’	/ɥà/
22.	‘potassium carbonate’	‘kanwa’	/sō/
23.	‘lion’	‘zaki’	/ɥ̥ā̤/
24.	‘elephant’	‘giwa’	/sō̤/
25.	‘vulture’	‘angulu’	/zà/
26.	‘rat’	‘ɓera’	/ʃə̤̀/
27.	‘dust’	‘kura’	/fō̤/
28.	‘island’	‘gungu’	/sà/
29.	‘foam’	‘kumfa’	/fō/
30.	‘slime’	‘yauki’	/pjə̤̀/
31.	‘mat’	‘zana’	/kī̤/
32.	‘knife’	‘wuka’	/ɥ̥à/
33.	‘slave’	‘bawa’	/ʃə̀/

Table 2: Minimal Triplets in the Wordlist:

S/No	**/CV/**		**/C̥V/**		**/C̥V̤/**	
	Word	*Gloss*	*Word*	*Gloss*	*Word*	*Gloss*
1	/də̀/	‘to hurry’	/tə̀/	‘to be tight’	/tə̤̀/	‘to give’
2	/dè/	‘to destroy’	/tè/	‘to spill’	/tè̤/	‘k.o grass’
3	/bə̀/	‘to see’	[1]/ǹ̩pə̀/	‘gamba grass’	/pə̤̀/	‘to pierce’
4	/zò/	‘to protrude’	/sò/	‘potassium carbonate’	/sò̤/	‘elephant’
5	/zà/	‘vulture’	/sà/	‘island’	/sà̤/	‘to read’
6	/bē/	‘to be lazy’	/pē/	‘to pass’	/pē̤/	‘to winnow’
7	/ɥà/	‘soup’	/ɥ̥à/	‘knife’	/ɥ̥à̤/	‘lion’

Table 3: **Exclusively minimal pairs**:

[1] In Dza, there are words that tend to carry an underlying [n̩] before the onset which could be realized as [m̩, ɲ̩, ŋ̩] depending on the place of articulation of the onset. I have not yet figured out the function of the syllabic nasal. However, it does not change the meaning of the word. This is common when words are said in isolation, but the syllabic nasal tend to dissappear in continues speech.

S/No	/CV/		/C̥V/		/C̥V̤/	
	Word	*Gloss*	*Word*	*Gloss*	*Word*	*Gloss*
1	-	-	/kù/	'to sieve'	/kṳ̀/	'to belch'
2	-	-	/pjə̀/	'to exchange'	/pjə̤̀/	'slime'
3	-	-	/kwə̀/	'to seize'	/kwə̤̀/	'to answer'
4	-	-	/ʃə̀/	'slave'	/ʃə̤̀/	'rat'
5	-	-	/kì/	'to cover'	/kḭ̀/	'mat'
6	-	-	/fō/	'foam'	/fò̤/	'dust'