Vocal characteristics of English-Northern Sotho bilingual:

a comparative study

by

Amy Hammann

u16032366

A dissertation submitted in fulfilment of the requirements for the degree MA Speech-Language Pathology in the Department of Speech-Language Pathology and Audiology, Faculty of Humanities, University of Pretoria

Supervisors:

Prof. Jeannie van der Linde

Mrs. Bhavani Pillay

December 2022

Amy Hammann

Department of Speech-Language Pathology and Audiology

University of Pretoria

Email: amyhammann97@gmail.com

UNIVERSITY OF PRETORIA

FACULTY OF HUMANITIES

DEPARTMENT OF SPEECH-LANGUAGE PATHOLOGY AND AUDIOLOGY

DECLARATION

Full Name: Amy Hammann

Student Number: 16032366

Degree: MA Speech-Language Pathology

I declare that this research report is my own original work. Where secondary material is used, this has been carefully acknowledged and referenced in accordance with university requirements.

I understand what plagiarism is and am aware of the University of Pretoria's policy in this regard.

27 November 2022

SIGNATURE **DATE**

Name of Principal Researcher(s)	Signature	Date
Amy Hammann		27/11/2022

Name of Supervisor(s)	Signature	Date
Bhavani Pillay		05/12/2022
Jeannie van der Linde		05/12/2022

Name of Head of Department	Signature	Date
Jeannie van der Linde		05/12/2022

ETHICS STATEMENT

The author, whose name appears on the title page of this dissertation, has obtained, for the research described in this work, the applicable research ethics approval.

The author declares that she has observed the ethical standards required in terms of the University of Pretoria's code of ethics for researchers and the Policy guidelines for responsible research.

ACKNOWLEDGEMENTS

- Thank you to my dad, Nicky Hammann, for motivating me and believing in me when I did not believe in myself. Thank you to my mom, Cindy Hammann, for reminding me to balance working and studying with fun during the stressful times.

- Thank you to my supervisors and mentors, Mrs Bhavani Pillay and Prof. Jeannie van der Linde, for guiding me to improve the article and dissertation with every submission. I am fortunate to have had such insightful and reliable supervisors.

- Thank you to Prof. Marien Graham for patiently explaining the data analysis process, and for conducting the statistical analysis.

- Thank you to Prof. Elsabe Taljard for kindly reviewing the translated Northern Sotho Rainbow Passage.

- Thank you to all the individuals who gave of their time to participate in the study.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

F_0: Fundamental frequency

KW: Kruskal-Wallis

PCC: Person-centred care

SLT: Speech-language therapist

SLP: Speech-language pathologist

WSR: Wilcoxon signed-rank

LIST OF TABLES

TERMINOLOGY

Acoustic analysis: The study of sound characteristics of speech involving physical aspects of spoken language (Nasser et al., 2006).

Bilingualism: The ability to apply two languages across a range of linguistic tasks, including speaking, reading, and writing, in a variety of contexts (Roberts et al., 2007).

Cross-linguistic effect: The interaction of two languages in a bilingual speaker as evidenced in the receptive and expressive use of language and other language-related actions, like gestures (Jarvis, 2012).

Fundamental frequency: The rate of vocal fold vibration that correlates perceptually with pitch (Manwa et al., 2012)

Home language: An individual acquires this language in their early years. This language is typically the individual's most natural instrument of thought and communication (Department of Basic Education, 2010).

Intensity: The energy of a sound with the perceptual correlate of loudness (Schwartz, 2004).

Language of learning and teaching: The language in which teaching, learning, and assessment occurs in schools.

Model voice: The acoustic parameters and perceptual quality of voice represents the speaker, in terms of gender, age, and emotional state, without undue effort or discomfort (Bradley, 2010).

Perceptual analysis: Describing the characteristics of voice using auditory perception (Ferrand, 2012).

Person-centred care: A philosophy of practice that encompasses an eclectic approach and partnership between the client and the clinician (DiLollo et al., 2010).

Sequential bilingual: An individual who acquires a second language after their first language has developed (Kroll et al., 2010).

Speech rate: A measurement of syllables per second relating to the perceived speed of speech production (Ferrand, 2012).

Speech-language therapist: The South African equivalent term for a speech-language pathologist (SLP). ASHA (2015) defines a SLP as a health care professional who works to prevent, assess, diagnose, and treat speech, language, voice, social communication, cognitive-communication, and swallowing disorders in children and adults. SLT is synonymous with SLP, and the use of either term depends on contextual preference. In Chapters 1, 2, and 4, SLT is used. In Chapter 3, SLP is used in line with journal requirements.

ABSTRACT

Introduction: Bilinguals constitute a significant portion of speech-language therapists' (SLTs) caseloads. Insight into the cross-linguistic effect on voice is needed to guide SLTs to make linguistically-appropriate observations when working with heterogenous populations.

Method: A comparative within-subject design explored vocal characteristics across 114 audio recordings. Nineteen female English-Northern Sotho bilingual individuals performed three speech tasks (reading, picture description and monologue) in each language. Acoustic analysis of mean fundamental frequency (f_0), intensity, and rate was conducted with Praat software. A blinded listeners' panel reached consensus during perceptual analysis.

Results: Across languages, mean f_0 and intensity were significantly greater during the Northern Sotho picture description ($p = 0.002$) and reading passage ($p = 0.033$) respectively. The English reading passage elicited a significantly quicker speech rate ($p = 0.002$). Within English, reading elicited a significantly quicker speech rate than the picture description ($p = 0.003$) and monologue ($p = 0.003$). Within Northern Sotho, reading elicited a significantly higher mean f_0 than the monologue ($p = 0.028$). Perceptual voice quality, glottal attack, and resonance were comparable across languages.

Discussion/Conclusion: Language, task performance, and vocal characteristics are not mutually exclusive in bilinguals. SLTs must consider the interaction between language, speech task, and voice when working with bilingual voice clients.

CHAPTER 1: Introduction

Chapter Aim: The ubiquitous nature of bilingualism is juxtaposed with the flexibility of the human voice to contextualise the present study. The current body of literature is reviewed. The rationale for the study is underpinned by the desideratum of South African SLTs working in communication focused contexts with diverse populations.

1.1 Background to the Study

Voice is a complex mechanism and the production thereof depends on a variety of factors (Ferrand, 2012). The physiology of voice production is thought to be universal, yet flexible (Nevo et al., 2015; Pépiot et al., 2021). The flexibility of the human voice affords individuals volitional and non-volitional modulations to reveal their intentions and index their identities (Lavan et al., 2018; Nevo et al., 2015; Pépiot, 2015). Ample research supports the linear connection between internal factors, such as age, gender, anatomy, and emotional state, and voice (Ferrand, 2012; Godin et al., 2015; Pépiot, 2014). However, very few studies have explored the dynamic relationships between external factors, such as language and speech tasks, when striving to understand voice in its entirety (Lee et al., 2017; Nevo et al., 2015; Pépiot et al., 2021).

Global interest in language variation continues to rise as bilingualism characterises contemporary society (Ansaldo & Saidi, 2014). Bilingualism is often considered a consequential life experience and the ability to communicate in more than one language is generally deemed an asset (Kroll et al., 2013). Fifty to eighty percent of the world is bilingual, and most of Africa's population is thought to be multilingual (Khokhlova, 2015; van Zyl et al., 2018). It is estimated that 2 110 languages are spoken across the continent (Bokamba, 2018; Lewis, 2009). However, the vocal characteristics of African languages remain unexplored. South Africa is a conducive

context in which to investigate the language-effect on voice as eleven official languages, including English and Northern Sotho, are recognised (Statistics South Africa, 2012).

Heterogeneity is densely represented in South Africa's bilingual populace as individuals present with an array of possible language pairs (Penn et al., 2017; Posel et al., 2016). English and Northern Sotho are widely spoken across South Africa (Statistics SA, 2012). Proficiency in both languages is desirable due to Northern Sotho's prevalence across the country and English's role in education and employment (Khokhlova, 2015; Statistics SA, 2012). The ubiquity of English-Northern Sotho bilingualism has not been statistically quantified. However, many South Africans are likely to communicate in both languages.

SLTs require a holistic understanding of the complex nature of voice to best-serve diverse populations; including English-Northern Sotho bilingual speakers (van der Merwe et al., 2014). Bilingual speakers acquire the morphology, syntax, semantics, phonology, and pragmatics, of two languages (Pineda-Pérez et al., 2021). Context, experiences, and non-linguistic components, like gestures and facial expressions, moderate language proficiency (Kroll et al., 2013; Pineda-Pérez et al., 2021). Bilingual speakers may inadvertently adjust their voice use when acquiring the pragmatic and non-linguistic aspects of their two languages which warrants further consideration (Pineda-Pérez et al., 2021).

Voice production, language, speech, and cognition differ across monolingual and bilingual populations as these groups do not process linguistic information in the same manner (Lee et al., 2017; Levi, 2018; Sadat et al., 2016). In comparison with their monolingual counterparts, bilingual speakers may take longer to articulate, make more

errors when speaking, and are more likely to experience the tip of the tongue phenomenon (Sadat et al., 2016). On the other hand, bilingual speakers may have unique advantages in intellectual development, creativity, flexibility, efficacy in shifting attention between tasks and openness to cultural diversity (Bialystok et al., 2004; Levi, 2018). Bilingual speakers are equipped with more robust linguistic processing, greater knowledge of linguistic structure, and improved auditory control (Levi, 2018; Theodore et al., 2020). These strengths may enable a bilingual advantage in voice processing and precipitate differing vocal characteristics across languages which, however, remain relatively unexplored (Theodore et al., 2020).

The human voice has sustained interest through decades of research (Cataldo et al., 2020). An individual's voice plays a fundamental role in transmitting knowledge and emotions across all contexts (Cataldo et al., 2020). Typical voice production occurs as the respiration, phonation, and resonance subsystems interact (Ferrand, 2012). During respiration, air is exhaled from the lungs to provide energy for sufficient subglottal pressure during phonation (Schwartz, 2004). The larynx is the site of phonation and comprises interlinked cartilages, membranes, ligaments, muscles, and soft tissue (Ferrand, 2012). During phonation, the vocal folds adduct and abduct symmetrically to self-sustain vibration and modulate glottal airflow while producing sound (Zhang, 2016). The degree of activation of intrinsic and extrinsic muscles of the larynx adapts the medial surface shape, tension, and approximation of the vocal folds. The sound stream is manipulated as it passes through the oral and nasal cavities during resonance and articulation (Schwartz, 2004).

The vocal apparatus is thought to adopt habitual language-specific configurations that are applicable at every level of phonetic analysis (Mennen et al., 2010). These

language-specific configurations have been termed phonetic settings (Honikman, 1964). Languages have different phonetic settings to accommodate their different phoneme inventories, phonation types, pitch ranges, register, and prosody (Mennen et al., 2010). Differing vocal characteristics may manifest as a function of varied phonetic settings across languages. However, the extent to which language exclusively affects voice has received sporadic interest and remains unclear in recent research (Pineda-Pérez et al., 2021).

1.2 Literature Overview

The language-effect on voice is intriguing as language is central to individual identity, and vocal characteristics can influence the listener's perceptions of the speaker (Altenberg et al., 2006; Mennen et al., 2010). In spontaneous speech, listeners attend to vocal characteristics that include f_0 range, speech rate and intensity (Kreiman & Sidtis, 2011). Differences across these suprasegmental aspects of speech may influence the meaning of the message as well as the speaker's perceived community membership, personality, competence, and attractiveness (Altenberg et al., 2006; Mennen et al., 2010). These parameters of voice are typically evaluated in a voice assessment (Ferrand, 2012). A voice assessment usually requires the client to perform a variety of speech tasks; irrespective of whether they are a monolingual or multilingual speaker (Ferrand, 2012). The intricate relationship between language, speech tasks, and voice must be investigated to support SLTs to understand culturally- and linguistically-based communication differences when working with bilingual speakers (Pépiot et al., 2021).

Early studies attempted to isolate the language-effect on voice across monolingual groups. English and German monolingual speakers produced significantly different f_0

ranges in their respective languages when reading a list of words, short phrases, and a short story (Mennen et al., 2012). Mandarin speakers presented with higher f_0 maximums, higher f_0 means, and wider f_0 ranges than their English counterparts in single word utterances (Keating et al., 2012). When saying CVCV pseudowords, French females' mean voice onset time (VOT) was longer in voiceless plosives, whereas English females had a longer mean VOT in aspirated plosives (Pépiot, 2015). There appeared to be a language- and task-effect on voice across monolingual groups, which sustained interest in the cross-linguistic effect on voice.

However, investigating the language-effect on voice across monolingual groups did not account for innate variables within each speaker that are known to influence voice (Awan et al., 1996; Nevo et al., 2015). It has been seen valuable to consider the language- and task-effect on voice in bilingual speakers in recent studies (Pineda-Pérez et al., 2021). Innate confounding variables, such as vocal tract anatomy and personality traits, are controlled as each bilingual speaker serves as their own control (Nevo et al., 2015).

English has been the most included language across bilingual dyads; owing to its global dominance across education and business (Pineda-Pérez et al., 2021; Posel et al., 2020). F_0 has been the most frequently considered parameter of voice as it is a salient feature of talker identity (Levi, 2018; Pineda-Pérez, 2021). F_0 depends largely on the instantaneous configuration of the larynx (Ng et al., 2012). Language-specific phonetic settings may manifest in f_0 values and give insight into the language effect on voice.

It is valuable to evaluate vocal characteristics across language pairs of differing origins to obtain robust results. Interest in Germanic and tonal language pairs has gained

momentum in cross-linguistic research. Vocal characteristics have been compared across English and: Japanese, Hebrew, Mandarin, Korean, and Cantonese (Graham, 2014; Lee et al., 2017; Ng et al., 2012).

Significant f_0 differences were identified when English-Japanese, English-Welsh, and English-Korean bilingual speakers read in their respective languages (Graham, 2014; Lee et al., 2017; Ordin et al., 2017). Reading in Japanese evoked higher pitch levels and wider f_0 spans in male and female English-Japanese bilingual speakers (Graham, 2014). The Welsh reading passage elicited wider f_0 span and higher maximum f_0 from female English-Welsh bilinguals in comparison with English (Ordin et al., 2017). Female English-Korean speakers presented with significantly higher f_0 when reading in Korean when compared to English reading tasks (Lee et al., 2017).

The task-effect on voice is not novel (Lee et al., 2017). Some speech tasks, such as sustained vowel phonation and tone sweeps, rely solely on the physiology of voice and, as such, do not functionally represent voicing for speech (Altenberg et al., 2006; Keating et al., 2012). These speech tasks could be independent from, and subsequently misrepresent, the language effect on voice. It is essential to consider the language effect on voice across connected speech tasks (Keating et al., 2012; Pineda-Pérez, 2021).

Read speech is distinguished from spontaneous speech by differing speech rates, mean f_0 and f_0 range (Hazan et al., 2010; Lavan et al., 2018). The structured, task-oriented nature of reading may cause speakers to hyperarticulate and present with higher f_0 measurements (Hudson et al, 1982; Lee et al., 2017). However, this consideration neglects language as a confounding derivative. Spontaneous speech is unstructured and reflects habituated, natural voice (Nevo et al., 2015). Explicit cross-

language f_0 variations in spontaneous speech suggest that languages do not manipulate the vocal apparatus in a universal manner (Cantor-Cutiva et al., 2021; Cheng, 2020).

In spontaneous speech, female English-Hebrew bilinguals presented with significantly higher mean f_0 in Hebrew. Similarly, male English-Hebrew bilinguals presented with significantly lower mean f_0 in English spontaneous speech (Nevo et al., 2015). Female English-Cantonese bilingual speakers had significantly higher f_0 values in English connected speech (Ng et al., 2012). Bilingual speakers may employ language-specific laryngeal configurations that induce f_0 variations across speech tasks (Pépiot et al., 2021).

Language-specific, predictable phonologic patterns may contribute to f_0 variations across speech tasks (Graham, 2014). The unique speech motor behaviour in tonal languages, such as tone variation, aspiration, and devoicing of speech sounds are distinguished from Germanic languages (Pinho et al., 2012). Lexical tones characterise tonal languages and may represent differing f_0 features that induce higher mean f_0 and wider f_0 ranges (Barcroft et al., 2014; Lee et al., 2017). Speaking a tonal language increases demands on neuromotor control (van der Merwe et al., 2014). Tonal languages require adept manipulation of the vocal folds, and an intricate intra-aural feedback system, to index meaning through lexical tones (Graham, 2014; Theodore et al., 2020). Tonal languages uniquely use syllable-level pitch variations to index semantic differences in phonologically similar words (Wong et al., 2009). For example, in Northern Sotho, lápà (high-low tone) means “tired” in contrast with làpá (low-high tone) which means “courtyard” (Malan, 2016). These unique intrinsic characteristics of tonal languages are expected to induce f_0 differences between

English, a Germanic language, and Northern Sotho, a tonal language, in the current study.

F_0 variability has been observed across Germanic-tonal pairs. However, the direction of change has been erratic (Cheng, 2020; Lee et al., 2017). The inconsistent results suggests that f_0 variations alone do not provide sufficient insight into the language effect on voice. Recent studies have revealed anecdotal evidence that listeners perceive language-specific vocal differences across speech rate, intensity, and perceptual parameters (Mennen et al., 2010; Nevo et al., 2015).

English-Mandarin and English-Korean bilingual speakers yielded comparable speech rates in English but spoke significantly faster in Mandarin and Korean (Lee et al., 2017). No task- or language-specific effect contributed to changes in intensity across English, Mandarin, and Korean (Lee et al., 2017). Nevo et al (2015) found that English-Hebrew males presented with hard glottal attack and vocal fry when speaking in English. However, English-Hebrew females presented with increased throaty resonance when speaking in Hebrew (Nevo et al., 2015). It is valuable for SLTs to take cognisance of language-and gender-based voice differences that may be perceptually salient. However, perceptual language-based voice differences have rarely been considered in bilingual voice studies (Nevo et al., 2015).

1.3 Rationale

South African SLTs work in unique communication-focused contexts with diverse populations (Mophosho, 2018). Interest in the English-Northern Sotho dyad is warranted as Northern Sotho is one of the three most spoken languages nationally. English remains influential as the dominant language in education, trading, public service, and government (Khokhlova, 2015; Statistics SA, 2012). English proficiency

is viewed as an avenue to socio-economic mobility as it is highly regarded in the South African labour market (Posel et al., 2020; Rudwick, 2008). English is a minority first language yet a common second language, as many South Africans acquire English at school and continue using it in the workplace (Posel et al., 2020). Fortunately, the widening of linguistic repertoires is more apparent than linguistic assimilation to English in South Africa (Deumert, 2010; Posel et al., 2020). Communicating in African languages remains desirable as it signifies ethnic identity, respect for African culture, and rejection of South Africa's colonial past (Posel et al., 2020). Many South Africans are presumed to be bilingual; owing to the need to preserve the African languages, although English is a dominant language in society (Posel et al., 2020).

Researchers have investigated the English-Northern Sotho language pair in literacy development, stroke rehabilitation, neuromotor speech disorders and sound system characteristics (Fouche et al., 1999; van der Merwe et al., 2014; van Zyl et al., 2018; Wilsenach et al., 2018). However, the vocal characteristics of English-Northern Sotho bilingual speakers remain unexplored. The English-Northern Sotho language pair is pertinent on a national level and novel on an international level (Statistics SA, 2012). Investigating a novel language pair is expected to make an eclectic contribution to the growing body of research interested in isolating language as an acquired influence on voice.

Insight into the language-effect on voice is needed to inform service delivery and augment person-centred care (PCC). For example, it may be necessary for SLTs to assess and treat bilingual voice clients in both languages. Counselling bilingual voice clients on the language-effect on voice may contribute to PCC. PCC can improve perceived quality of therapy, client satisfaction, and ultimately therapy outcomes

(DiLollo et al., 2010). SLTs could consider the interaction between languages and speech tasks a fundamental component of PCC as SLTs work in unique communication-focused contexts (Mophosho, 2018).

1.4 Concluding Statement

The fluid presentation of the language-effect on voice has sustained interest in isolating language as an acquired influence on voice (Cheng, 2020; Lee et al., 2017; Nevo et al., 2015; Ng et al., 2012). It is essential to investigate differing vocal characteristics across languages to support SLTs in making linguistically appropriate observations when working with heterogenous populations (Altenberg et al., 2006; Mophosho, 2018). The following question is posed: How do vocal characteristics compare across English and Northern Sotho speech tasks?

CHAPTER 2: Method

Chapter Aim: The research aim, design, ethical considerations, inclusion criteria, equipment, procedures, and considerations for reliability and validity are delineated to provide a detailed explanation of the research process.

2.1 Research Aim

The study aimed to investigate the vocal characteristics of English-Northern Sotho bilingual females across languages and speech tasks that included: a reading passage, picture description, and monologue.

2.2 Research Design

The study followed a comparative within-subject design to investigate multiparametric factors in a controlled environment (Lee et al., 2017; Nevo et al., 2015). The comparative approach allowed the relationship between language, voice, and task performance to be quantified. Within-subject comparisons were made as each participant served as their own control (Jackson, 2012; Vogt et al., 2012).

2.3 Ethical Considerations

Ethical research necessitates good conduct towards participants, colleagues, and the broader society while conducting a study (Vogt et al., 2012). Ethical clearance was obtained from the Research and Ethics Committee, Faculty of Humanities (HUM018/0221) (Appendix A). The following considerations informed ethical practice throughout this study:

2.3.1 Voluntary Participation

Prospective participants were not obligated to participate. Before signing the consent form, each participant was informed that they were free to terminate their participation in the study at any time, with no penalties (Leedy et al., 2016).

2.3.2 Informed Consent

Informed consent protects the participant's right to autonomy (Leedy et al., 2016). The nature of the study was explained verbally and presented in writing. Each participant had the opportunity to ask questions about the study. The participants read and signed an informed consent letter (Appendix B) that detailed: the purpose of the study, the requisites of the participants, the risks and benefits involved in participating, and the intended use of the data obtained.

2.3.3 Protection from Harm

The risks of participating in a research study should not exceed the risks involved in one's everyday living (Leedy et al., 2016). No risks for potential harm were identified for this study's participants. Data was collected during the COVID-19 pandemic. Measures, such as wearing masks, sanitising, and social distancing, were undertaken to minimise the spread of COVID-19.

2.3.4 Right to Privacy

The researcher maintained the participants' right to privacy throughout the study (Leedy et al., 2016). An alpha-numeric code was assigned to each participant in lieu of their names. The alpha-numeric code was used during data collection, data

processing, data storage, and reporting of results, for example, P01 = participant 1. In the event of publication, no personally identifying information will be disclosed.

2.3.5 Honesty

Information and results were truthfully reported. The procedures were not misrepresented and the participants were not misled (Leedy et al., 2016). The publications and resources that contributed to this study were acknowledged and referenced according to the American Psychological Association (7th ed.) referencing style in the dissertation. The article included Vancouver style referencing, in line with the journal requirements.

2.3.6 Data Storage

Data will be published and managed in an institutional Research Data Management system, an accredited open data repository, or an accredited discipline-specific repository, in accordance with the University of Pretoria's Research Data Management Policy.

2.4 Participants

2.4.1 Sampling

The snowball sampling technique allowed the researcher to identify participants in a time- and cost-effective manner. The advertisement (Appendix C) was posted on social media. The advertisement detailed the: site of the study, inclusion criteria, and contact details of the researcher. Interested viewers contacted the researcher to volunteer to participate. Prospective participants shared the advertisement and

encouraged potential participants known to them to contact the researcher to volunteer to participate.

- ***Female.*** The most conspicuous difference between male and female speakers is f_0. Females speak at an average of 210Hz and males speak at an average of 120Hz (Ferrand, 2012). It was valuable to include participants of one gender only to eliminate the confounding influence of gender on voice.

- ***English-Northern Sotho Bilingual Speakers.*** Including bilingual speakers allowed the researchers to better isolate the language effect on voice as intrinsic confounding variables were controlled. Proficiency in English and Northern-Sotho was determined based on the participants' responses in section three of the participant questionnaire (Appendix D). Participants were required to (a) consider themselves bilingual and (b) report on their confidence communicating in both languages across different contexts with a variety of communication partners. Valid within-subject comparisons were to be made as each participant served as their own control.

- ***Age: 18 – 65 years old.*** Differences in auditory-perceptual, visual, and acoustic findings indicate that the voice changes as the speaker ages (Ferrand, 2012). The voice stabilises after puberty once the larynx has developed its adult shape, size, and structure (Ferrand, 2012). The ageing mechanism of the larynx has been termed presbyphonia (Ferrand, 2012). It was viable to include participants aged 18 to 65 as their voices would have reached maturation without being compromised by presbyphonia. ***Normal Hearing.*** People with hearing loss have difficulty monitoring their vocal output due to their lack of auditory feedback caused by the hearing loss (Coelho et al., 2015). The type and

severity of hearing loss determine the accompanying voice problems (Coelho et al., 2015). Voice problems may present as: strain, breathiness, roughness, monotone, hoarseness, reduced pitch, and reduced volume (Coelho et al., 2015). Participants with normal hearing were included in the study to minimise the extraneous effect of hearing loss on vocal characteristics to improve the internal validity of the research.

2.4.1.2 Exclusion Criteria. Prospective participants were excluded from the study if they had been diagnosed with a voice disorder or failed the hearing screening.

2.4.1.3 Participant Description. All nineteen participants were female. Their ages ranged from 18 years old to 54 years old. Normal hearing was confirmed, and normal vision was self-reported. All the participants considered themselves bilingual with healthy voices. They were comfortable communicating in both English and Northern Sotho across contexts. Additional information about the participants' acquisition and daily use of English and Northern Sotho is presented in Table 1.

Table 1

Participants' language acquisition and use information

Acquisition and daily use	English		Northern Sotho	
	n	%	n	%
Period of acquisition				
First language	0	0	17	89.5
Preschool	11	57.9	1	5.3
Primary school	8	42.1	1	5.4
Daily environmental use				
Home	6	31.6	18	94.7
Work	9	47.4	10	52.6

Social	14	73.7	13	68.4
Daily communication partners				
Family	6	31.6	17	89.5
Co-workers	15	78.9	8	42.1
Friends	17	89.5	13	68.4

2.5. Materials and Apparatus

The equipment needed to conduct the study is presented in Table 2.

Table 2

Materials and Apparatus

Component	Material / Apparatus	Description
Participant screening	hearScreen application	The hearScreen application provides clinically valid pure tone audiometric screening results in a timeous manner (Louw et al., 2017).
	Smartphone	The hearScreen application was downloaded onto a Samsung Galaxy J5 Prime smartphone.
	Headphones	Supra-aural headphones were calibrated according to ISO/ANSI standards. The headphones were plugged in to the smartphone to conduct the hearing screening.
	Participant questionnaire	The researcher developed a questionnaire (Appendix D) based on the literature. The self-administered questionnaire comprised a series of dichotomous, closed, and open-ended questions. Voice, and language information was obtained from the participants.
Data Collection	Written Instructions	Written instructions corresponded with the verbal instructions presented before the speech task.

	Audio Recorder	A Philips VoiceTracer Audio Recorder (DVT1150) was used to record each participant performing the speech tasks.
	The Rainbow Passage	The Rainbow Passage is widely used in clinical practice and research (Fairbanks, 1960). The original Rainbow Passage (Appendix E.1) was translated into Northern Sotho and reviewed by a qualified linguist for this study (Appendix E.2).
	The Cookie Theft picture	The Cookie Theft picture (Appendix E.3) from the Boston Diagnostic Aphasia Examination (Goodglass et al., 1972) extracts a wide range of predictable abstract and concrete linguistic content from the speaker.
	Monologue Topics	Participants selected one of two topics for both monologues (Appendix E.4).
Acoustic Analysis	Laptop	Data was stored and processed on a Lenovo B50-80 Notebook using Audacity v2.3.2 software (Softonic , 2019), Praat, and Microsoft Excel.
	Praat software v5.4.56	Praat is freeware often used in clinical studies (Boersma et al., 2019). Praat software was used to conduct an acoustic analysis of mean f_0, intensity, and speech rate.
Perceptual Analysis	GRBASI scale	The GRBASI scale (Appendix F) reliably measures a speaker's perceptual voice quality (Yamauchi et al., 2010). The 4-point rating scale evaluates hoarseness, roughness, breathiness, asthenia, strain, and instability.
	Nominal Scale	Resonance and glottal attack were evaluated by replicating a nominal scale by Nevo et al. (2015) (Appendix F). The scale characterises glottal attack as hard, soft, or adequate. Resonance could be rated as adequate, nasal, or throaty.

2.6 Procedures

2.6.1 Data Collection and Recording

Data collection commenced once ethical clearance (Appendix A) was granted. The researcher posted the advertisement (Appendix C) across social media platforms. Viewers shared the advertisement to promote snowball sampling. Prospective participants contacted the researcher to enquire or volunteer to participate.

A one-hour appointment was scheduled with each prospective participant at the sound-attenuated Voice Laboratory (Room 3-19) in the Department of Speech-Language Pathology and Audiology, University of Pretoria. All participants were screened for symptoms of COVID-19. A social distance of 1 - 2 meters between the researcher and the participant was maintained throughout to minimise the risk of the spread of COVID-19. Alcohol-based sanitizer was readily available and used to sanitize surfaces and apparatus between sessions.

Written informed consent (Appendix B) was obtained from each participant. The participant completed the questionnaire (Appendix D) and the hearing screening was conducted. Instructions were presented in written and verbal format to each participant in English. Translating instructions into Northern Sotho was unnecessary as all participants were proficient in both languages. Each participant performed three different speech tasks (a reading passage, picture description, and monologue) in English and in Northern Sotho while being audio recorded. Participants were given brief breaks to combat fatigue between the six speech tasks.

The languages were counter-balanced, and the sequence of speech tasks was randomised across participants to control for order effects. The audio recorder was at

a constant distance of 45cm from the speaker on the table. High-quality recording and low microphone sensitivity settings were consistently applied to support the reliability of the study.

Each participant needed to remove their mask while performing the speech tasks to preserve the integrity of the data. All participants were comfortable removing their masks, although they could decline participation in the study if they were not comfortable. This is consistent with voluntary participation as an ethical consideration.

The participants familiarised themselves with the Rainbow Passages (Appendices E.1 – E.2) and the Cookie Theft picture (Appendix E.3). Each participant was audio recorded while they read and described the picture in each language. The participants spoke about either of the monologue topics (Appendix E.4) for one minute. The same topic was used per participant across languages. A total of six audio recordings were obtained from each participant; three in English and three in Northern Sotho.

2.6.2 Data Processing and Analysis

The audio recordings were imported from the Philips VoiceTracer Audio Recorder (DVT1150) to the A Lenovo B50-80 Notebook. The recordings were saved in .WAV format and imported to Audacity v2.3.2, where silent pauses exceeding two seconds were omitted.

2.6.2.1 Acoustic Analysis

Acoustic analysis provides quantitative information about voice production and is an invaluable component of a voice assessment (Ferrand, 2012). Praat was used to acoustically analyse all 119 audio recordings in terms of: f_0, intensity and speech rate.

These acoustic parameters were selected for this study to yield results comparable to the existing body of research (Lee et al., 2017; Mennen et al., 2012; Nevo et al., 2015).

The recordings were exported to Praat in ".WAV" format. The researcher clicked on the recording, listed as an "Object" in Praat and selected "View and Edit". Mean f_0 and intensity were obtained by highlighting the entire speech sample and selecting "Pitch" and "Intensity" to "Get Pitch" and "Get Intensity", respectively, on the toolbar.

A Praat script that was validated by de Jong et al (2009) measured the speech rate. The script detects syllable nuclei and provides an automatic measure of speech rate in syllables per second (de Jong et al., 2009). The script was inserted as a "New Script". Settings were not adjusted, but the directory was changed to the location of the audio recording as saved on the laptop's hard drive. The script was applied, and the speech rate was obtained. The mean f_0, intensity, and speech rate of each recording were copied and pasted into the Excel spreadsheet containing the raw data.

2.6.2.2 Perceptual Analysis

The GRBASI and nominal scales (Appendix F) were used by the blinded listeners' panel to evaluate each speaker's perceptual voice quality, resonance, and glottal attack in all 38 monologues. The blinded listeners' panel comprised three SLPs, all of whom were familiar with English and Northern Sotho. A blinded listeners' panel increased the reliability and validity of the perceptual analysis. The monologue was selected for perceptual analysis since it best represented natural communication (Kreiman et al., 2011). The panel listened to each monologue recording via the Lenovo B50-80 Notebook. The order of languages was randomised to control for order effects during the perceptual analysis. The panel members completed the GRBASI and

nominal scales (Appendix F) and discussed their results to reach a consensus for each recording.

2.6.2.3 Statistical Analysis

Descriptive and inferential statistics were conducted using the Statistical Package for the Social Sciences (SPSS v.26). Non-parametric tests were deemed appropriate due to the small sample size ($n < 30$) and the lack of normal distribution across the continuous variables. The Wilcoxon signed-rank (WSR) test was used to analyse the continuous acoustic variables, including: f_0, intensity, and speech rate. The Kruskal-Wallis (KW) test compared the acoustic parameters within English and Northern Sotho. Dunn's post-hoc pairwise testing was conducted when significant, task-based acoustic differences within either language were noted. The categorical perceptual variables obtained from the GRBASI and nominal scales were analysed using McNemar's test (for 2x2 crosstabulations) and the test for marginal homogeneity (for crosstabulations greater than 2x2).

2.7 Reliability

Reliability refers to the consistency of a measurement (Kaur et al., 2018). Environmental errors were minimised as all recordings were captured in the same sound-attenuated booth. The audio recorder was always positioned at the same distance from the speaker, and consistent recording and sensitivity settings were applied. The order of speech tasks was randomised, and languages were counter balanced to control for order effects. Participant changes were expected to be insignificant as all audio recordings were captured on the same day, within one hour from each participant. The participants took brief breaks between speech tasks to combat fatigue.

2.8 Validity

Validity entails the extent to which the scores from a measurement represent the intended variable (Kaur et al., 2018). The validity of the study was preserved as the selected elicitation material, and analyses software have been widely used in previous research. The Rainbow Passage (Fairbanks, 1960), Cookie Theft Picture (Goodglass et al., 1972), GRBASI Scale (Yamauchi et al., 2010), and Praat (Boersma et al., 2019) are standardised and validated tools.

CHAPTER 3: Research Article

Chapter Aim: This chapter introduces the journal selected for publication and outlines the journal's submission guidelines. The submitted article succeeds this information, as per the departmental postgraduate format, and is formatted according to the journal specifications.

The International Journal of Phoniatrics, Speech Therapy and Communication Pathology has been published since 1949. The journal publishes international research on the anatomy, physiology, and pathology of the structure of speech, language, swallowing, and hearing mechanisms. Papers may report new findings on basic function, assessment, management, and test development. Specific theories of speech language, swallowing, and hearing function may be presented.

Title: Vocal characteristics across English-Northern Sotho bilingual individuals: a comparative study

Authors: Amy Hammann, Bhavani Pillay, Marien Alet Graham, and Jeannie van der Linde

Journal: International Journal of Phoniatrics, Speech Therapy and Communication Pathology

Submitted for review: 27th of November 2022 (Appendix G)

Note: This article was compiled according to the editorial specifications of this journal, and subsequently differs from the editorial style of the rest of the dissertation.

Vocal characteristics across speech tasks in English-Northern Sotho bilinguals: a comparative study

Abstract

Introduction: Bilinguals constitute a significant portion of speech-language pathologists' (SLPs) caseloads. Insight into the cross-linguistic effect on voice is needed to guide SLPs to make linguistically-appropriate observations when working with heterogenous populations.

Method: A comparative within-subject design explored vocal characteristics across 114 audio recordings. Nineteen female English-Northern Sotho bilingual individuals performed three speech tasks (reading, picture description and monologue) in each language. Acoustic analysis of mean fundamental frequency (f_0), intensity, and rate was conducted with Praat. A blinded listeners' panel reached consensus during perceptual analysis.

Results: Across languages, mean f_0 and intensity were significantly greater during the Northern Sotho picture description ($p = 0.002$) and reading passage ($p = 0.033$) respectively. The English reading passage elicited a significantly quicker speech rate ($p = 0.002$). Within English, reading elicited a significantly quicker speech rate than the picture description ($p = 0.003$) and monologue ($p = 0.003$). Within Northern Sotho, reading elicited a significantly higher mean f_0 than the monologue ($p = 0.028$). Perceptual voice quality, glottal attack, and resonance were comparable across languages.

Discussion/Conclusion: Language, task performance, and vocal characteristics are not mutually exclusive in bilinguals. SLPs must consider the interaction thereof when working with bilingual voice clients.

Introduction

Speech-language pathologists (SLPs) require a holistic understanding of the complex nature of voice to best-serve the heterogenous population with whom they interact [1]. Ample research supports the linear connection between internal factors, such as age, gender, anatomy, emotional state, and voice [2, 3]. However, a small number of studies have explored the dynamic relationship between external factors, such as language and speech tasks, when striving to understand voice in its entirety [4 – 6].

The extent to which language exclusively affects voice has received sporadic interest and remains unclear in recent research [7]. It is essential to consider the language effect on voice as 50% to 80% of the global population is bilingual [8]. Linguistic diversity is densely represented in Africa. It is estimated that 2 110 languages are spoken across the continent [9; 10]. Africa is a conducive context in which to investigate the language effect on voice as most of the population is thought to be multilingual [11]. However, the vocal characteristics of African languages remain unexplored. Only a few studies have considered bilingual speakers' vocal characteristics across English and other languages including: Japanese, Hebrew, Welsh, Mandarin, Korean, and Cantonese [4, 5, 12 – 14]. Fewer studies recognised speech tasks as a confounding factor when attempting to isolate language as an acquired influence on voice [4 – 6].

Fundamental frequency (f_0) is a salient feature of talker identity and has been widely considered in cross-linguistic studies [15]. Significant f_0 differences were identified when English-Japanese, English-Welsh, and English-Korean bilingual speakers read in their respective languages [5, 12, 13]. Reading in Japanese evoked higher pitch levels and wider f_0 spans in male and female English-Japanese bilingual speakers [12]. The Welsh reading passage elicited wider f_0 span and higher maximum f_0 from female English-Welsh bilinguals

compared to English [13]. Female English-Korean speakers presented with significantly higher f_0 when reading in Korean as compared to English reading tasks [5].

The task effect on voice is not novel [5]. The structured nature of reading may cause speakers to hyperarticulate and present with higher f_0 measurements [5, 16]. However, this consideration neglects language as a confounding derivative. Spontaneous speech tasks reflect habituated natural voice [4]. Explicit cross-language f_0 variations in spontaneous speech suggest that languages do not manipulate the vocal apparatus in a universal manner [17, 18]. In spontaneous speech, female English-Hebrew bilinguals presented with significantly higher mean f_0 in Hebrew. Similarly, male English-Hebrew bilinguals presented with significantly lower mean f_0 in English spontaneous speech [4]. Female English-Cantonese bilingual speakers had significantly higher f_0 values in English connected speech [14]. Bilingual speakers may employ language-specific laryngeal configurations that induce f_0 variations across structured and unstructured speech tasks [6].

Language-specific, predictable phonologic patterns may contribute to f_0 variations [12]. Greater f_0 variability may be evident in tone and pitch-accent languages as the vocal folds are adeptly manipulated to index meaning through lexical tones [1, 12]. Varying tone, aspirating and devoicing speech sounds may distinguish tonal languages from Germanic languages [19]. F_0 variability has been observed across Germanic-tonal pairs, although the direction of change has been erratic [5, 18]. The inconsistent findings suggest that f_0 variations alone do not provide sufficient insight into the language effect on voice.

Recent studies revealed anecdotal evidence that listeners perceive language-specific vocal differences across speech rate, intensity, and perceptual parameters [4, 20]. English-Mandarin and English-Korean bilingual speakers yielded comparable speech rates in English but spoke significantly faster in Mandarin and Korean [5]. English spontaneous speech presented with

greater incidence of hard glottal attack and vocal fry in comparison with Hebrew connected speech [4]. No task- or language-specific effect contributed to changes in intensity across English, Mandarin, and Korean [5].

Vocal characteristics are contingent on the language and speech task, although little is known about the interaction thereof. Further multiparametric studies are needed to understand vocal characteristics as a function of language and speech tasks. The following question is posed: How do vocal characteristics compare in English-Northern Sotho bilingual speakers across speech tasks?

Materials & Method

A comparative within-subject design was selected to quantify English and Northern Sotho vocal characteristics across a reading passage, picture description and monologue.

Participants

Nineteen female English-Northern Sotho bilingual speakers volunteered to participate after viewing an advertisement about the study on social media platforms. The volunteers contributed to snowball sampling by suggesting prospective participants known to them. The inclusion criteria required participants to be: (1) female; (2) between 18 and 65 years old; (3) proficient in English and Northern Sotho; (4) presenting with normal hearing. The inclusion criteria controlled for confounding gender, age, language proficiency, and hearing loss variables [2, 21].

The hearScreen application confirmed that each participant had normal hearing [22]. Each participant completed a questionnaire detailing relevant voice and language information. The mean age of the participants was 22.5 (SD = 7.848) years old. All participants felt that they

had normal healthy voices. Participants denoted proficiency in English and Northern Sotho by (a) considering themselves bilingual and (b) reportedly speaking both languages across contexts with a variety of communication partners, as presented in Table 3.1. Including bilingual speakers allowed valid cross-linguistic comparisons to be made as each participant served as their own control.

Table 3.1 *Participants' language acquisition and use information*

Acquisition and daily use	English		Northern Sotho	
	n	%	n	%
Period of acquisition				
First language	0	0	17	89.5
Preschool	11	57.9	1	5.3
Primary school	8	42.1	1	5.4
Daily environmental use				
Home	6	31.6	18	94.7
Work	9	47.4	10	52.6
Social	14	73.7	13	68.4
Daily communication partners				
Family	6	31.6	17	89.5
Co-workers	15	78.9	8	42.1
Friends	17	89.5	13	68.4

Procedure

Procedures were similar to those of Lee et al., [5] and Nevo et al. [4], as outlined in Table 3.2.

Table 3.2 *Voice analysis protocol adapted from Lee et al., (2017) and Nevo et al. (2015)*

Analysis	Apparatus	Speech Task(s)	Parameters
Acoustic	Praat software (v.5.4.56) [32]	Reading passage Picture description Monologue	Mean f_0 Intensity Speech rate
Perceptual	GRBASI scale [33] Nominal scale	Monologue	Voice quality Resonance Glottal attack

A Phillips VoiceTracer Audio Recorder (DVT1150) with standard settings was used to obtain audio recordings in a sound-attenuated booth. Verbal and written instructions were provided and required each participant to speak in their most natural comfortable voice. Participants were encouraged to take brief breaks between recordings to combat fatigue.

Speech tasks. Each participant performed a reading passage, picture description and a monologue in English and in Northern Sotho. These speech tasks were selected to elicit semantically matched samples across languages. The languages were counterbalanced, and the sequence of speech tasks was randomised to control for order effects.

The Reading Passage. The Rainbow Passage has been used in cross-linguistic studies investigating the language effect on voice [5, 23, 24]. The original Rainbow Passage amounts to 330 words in 19 sentences in English. It was translated to Northern Sotho and reviewed by a qualified linguist for this study. The Northern Sotho version amounted to 376 words in 19 sentences. Participants had the opportunity to practice reading each passage 2-3 times before being audio recorded.

The Picture Description. The "Cookie Theft" image from the Boston Diagnostic Aphasia Examination [25] elicited predictable, abstract and concrete discourse from participants. Participants acquainted themselves with the image and described the picture for approximately one minute in each language.

***The Monologue*.** A spontaneous, connected speech sample was elicited by asking participants to speak about a neutral topic, either "my family" or, "why I love South Africa" in both languages, for approximately 1 minute each. The same topic was used per participant across languages.

Data Analysis

The acoustic analysis was conducted in line with Lee et al.,'s [5] method, and the perceptual analysis was similar to Nevo et al. [4]. Praat is freeware often used in clinical studies [5]. Praat software was used to conduct an acoustic analysis of mean f_0, intensity, and speech rate. The monologue was selected for perceptual analysis as it best represents natural communication [26]. A blinded listeners' panel of three SLPs conducted perceptual analysis and reached consensus. All the SLPs were familiar with English and Northern Sotho.

Descriptive and inferential statistics were conducted using the Statistical Package for the Social Sciences (SPSS v.26). Non-parametric tests were deemed appropriate due to the small sample size ($n < 30$) and the lack of normal distribution across the continuous variables. The Wilcoxon signed-rank (WSR) test was used to analyse the continuous acoustic variables including: f_0, intensity, and speech rate. The Kruskal-Wallis (KW) test compared the acoustic parameters within English and within Northern Sotho. Dunn's post-hoc pairwise testing was conducted when significant, task-based acoustic differences within either language were noted. The categorical perceptual variables obtained from the GRBASI and nominal scales

were analysed using McNemar's test (for 2x2 crosstabulations) and the test for marginal homogeneity (for crosstabulations greater than 2x2).

Results

The comparison of the voice characteristics in English-Northern Sotho bilingual speakers across speech tasks is presented according to the outcomes of the acoustic and perceptual analysis.

Acoustic Analysis

When comparing the English and Northern Sotho speech tasks, only the picture description elicited a significantly higher mean f_0 in Northern Sotho (WSR = -2.938, p=0.002). Participants read with significantly greater intensity in Northern Sotho than in English (WSR = -2.113, p = 0.033). However, the participants read significantly faster in English than in Northern Sotho (WSR = -2.918, p = 0.002). No statistically significant differences were identified in the monologue task, as evident in Table 3.3.

In English, the reading passage elicited a quicker speech rate than the picture description task (Dunn's test statistic = 15.868, p = 0.003) and the monologue (Dunn's test statistic = 15.868, p = 0.003). In Northern Sotho, the participants presented with significantly higher mean f_0 during the reading task than in the monologue (Dunn's test statistic = 11.842, p = 0.028). As shown in Table 3.4, pairwise comparisons for intensity were not conducted as intensity did not differ significantly across speech tasks within either language.

Perceptual analysis

McNemar's test and the test of marginal homogeneity were used to test for differences in the nominal data including: voice quality, resonance, and glottal attack. All p-values were greater than 0.05, so no significant differences were found.

Table 3.3 *Acoustic parameter comparisons across speech tasks for English and Northern Sotho*

Acoustic parameter	English	Northern Sotho	WSR	p-value
Fundamental Frequency	Mean (SD)	Mean (SD)		
Reading passage	199.570 (21.787)	203.834 (19.600)	-1.288	0.210
Picture description	196.500 (23.178)	204.608 (19.486)	-2.938	0.002*
Monologue	194.869 (24.830)	191.110 (21. 698)	-1.288	0.210
Overall	196.980 (22.302)	199. 851 (18.703)	-1.650	0.104
Intensity				
Reading passage	65.089 (4.142)	66.378 (4.279)	-2.113	0.033*
Picture description	65.014 (4.999)	65.884 (4.702)	-1.650	0.104
Monologue	64.377 (5.435)	63.787 (5.487)	-0.966	0.352
Overall	64.827 (4.571)	65.350 (4.350)	-1.154	0.265
Speech Rate				
Reading passage	3.799	3.403	-2.918	0.002*

	(0.301)	(0.513)		
Picture description	3.399	3.665	-1.771	0.080
	(0.516)	(0.421)		
Monologue	3.459	3.680	-1.852	0.065
	(0.297)	(0.413)		
Overall	3.553	3.583	-0.040	0.984
	(0.322)	(0.323)		

*Statistical significance ($p < 0.05$)

Table 3.4 *Task-based acoustic differences within English and Northern Sotho*

Language	Acoustic parameter	Pairwise comparison	Dunn's test statistic	p-value
English	Rate	description – reading	15.868	0.003*
		reading – monologue	15.868	0.003*
		monologue-description	0.000	1.000
Northern Sotho	F_0	description – reading	-0.789	0.883
		reading – monologue	11.842	0.028*
		monologue-description	12.632	0.019*

*Statistical significance ($p < 0.05$)

Discussion/Conclusion

The results of this study revealed significant language- and task- effects on mean f_0, intensity, and speech rate across and within English and Northern Sotho speech samples. Perceptual

voice quality, glottal attack, and resonance between the English and Northern Sotho monologues were comparable.

F_0 is the most widely considered parameter in cross-linguistic voice studies (7). F_0 depends on the instantaneous configuration of the larynx as mediated by the language's typical vocal tendency [14, 20]. There is convergence on the idea that intrinsic features of tonal languages precipitate higher mean f_0 [5, 12, 18]. Tonal languages present with increased lexical high tones in connected speech that could elevate f_0 measurements [1, 18, 27].

Lee et al., [5] and Cheng [18] found that English-Korean female bilinguals presented with higher mean f_0 during reading, picture description, and monologue tasks in Korean when compared with English. Similarly in the current study, Northern Sotho speech had a higher mean f_0, although the difference was insignificant when compared with English overall. The mean f_0 was significantly higher in the Northern Sotho picture description than in the English picture description. The picture description task required participants to use predictable language to construct original discourse. The fixed-but-flexible nature of the picture description task may have revealed f_0 differences across languages that did not manifest during the unstructured monologue and the structured reading passage.

Significant f_0 differences were not identified in the English speech tasks. However, the mean f_0 was significantly higher during the Northern Sotho reading task than during the Northern Sotho monologue. These findings suggest that f_0 variations do not manifest in the same manner across languages even when identical elicitation tasks are used. Clear f_0 contrasts within the Northern Sotho samples indicate that tonal languages may be more susceptible to f_0 variations in specific speech tasks than non-tonal languages; in this case, English. SLPs should take cognizance of possible language and task effects in practice by assessing and treating bilingual voice clients using multiple elicitation tasks in more than one language.

All the participants were sequential bilinguals, and the acquisition of Northern Sotho preceded their acquisition of English. Many South Africans acquire English at school and continue speaking a tonal language at home [28]. Acquiring English as an academic language necessitates reading regularly in English [29]. Reading aloud may underscore a language's acoustic properties as words are inadvertently hyperarticulated throughout the structured task [5, 18]. The nature of reading aloud, paired with the participants' familiarity with reading aloud in English, may have contributed to differing vocal variations during the English reading passage.

The English-Northern Sotho female bilinguals read significantly quicker in English than they did in Northern Sotho. This result contrasts with English-Mandarin and English-Korean bilingual females who read significantly slower in English [5]. Speech rate was hypothesised to be quicker in Northern Sotho as morphemes are articulated in isolation, whereas English comprises slow syllabic rate with high-density syllables [5, 28]. English was the participants' non-native language in both the current study and in Lee et al.,' [5] study. However, approximately 65% of South Africans are taught in English throughout primary school [31]. The quicker speech rate evoked by the English reading passage in the current study may be ascribed to the participants' familiarity with reading in English given the language's role in South African education.

The bilingual's daily language use is an essential context within which to evaluate the interaction between language and speech tasks. Most of the participants reported conversing in Northern Sotho at home with family. The stress associated with performing an unfamiliar speech task can elevate f_0 measurements [24]. Reading less frequently in Northern Sotho may have resulted in a significantly higher mean f_0 during the Northern Sotho reading passage than during the Northern-Sotho monologue. Mean f_0 differences within the Northern Sotho samples suggests that the task-effect might be more transparent in tonal languages than in

Germanic languages, such as English. It is valuable for SLPs to interpret language- and task-based vocal variations against the context of each bilingual client's daily speaking demands.

Intensity has seldom been compared across languages, although it communicates the speaker's emotional state and personality [5, 26]. The Northern Sotho reading passage elicited significantly louder speech than the English reading passage. No significant task effect on intensity was noted within the English samples and within the Northern Sotho samples; corroborating with Lee et al., [5] findings. This result suggests that the interaction between the language and speech task may moderate intensity to a greater extent than either language or task effects in isolation.

The monologue task was thought to best-represent natural communication [26]. No significant acoustic or perceptual differences manifested across the English and Northern Sotho monologues. Language-based acoustic fluctuations were most prevalent during the reading task. When the reading task was compared with the monologues within each language, it was found that intensity and mean f_0 differed uniquely in English and Northern Sotho respectively. The bilinguals read with a significantly quicker speech rate in English yet read with a significantly higher mean f_0 in Northern Sotho. Contrasting structured and unstructured speech tasks highlights the unique way speech tasks interacted with each of the bilinguals' languages. The findings indicate it is valuable to compare vocal characteristics across a range of structured and unstructured tasks when assessing bilingual speakers.

The results of the current study corroborate with previous work that suggests vocal characteristics are contingent on the interaction between languages and speech tasks in the bilingual populace [5, 7]. It is valuable for SLPs to gain insight into bilingual speakers' language profiles, and vocal characteristics across languages during voice assessments. Acoustic analysis may reveal vocal characteristics that are neither perceptually obvious to the

SLP nor evident in only one of the bilingual's languages. Future cross-linguistic voice studies should include male and female participants [13, 17]. Females are more likely than males to ascribe to societal expectations and internalise social practices that may manifest in the voice [12, 18]. Subsequently, vocal variations may be more marked across languages in female speakers. However, further robust cross-linguistic voice research is required to explore this notion.

In conclusion, the language spoken by an individual is considered an acquired influence on voice. Differing manifestations of acoustic voice parameters across and within English and Northern Sotho samples suggest that languages do not interact with speech tasks in a universal manner. SLPs should consider task performance in different languages across different tasks when working with bilingual individuals. Further insight into the cross-linguistic effect on voice is needed to inform clinical assessment and decision-making.

Acknowledgements

The authors extend appreciation to Prof. Elsabé Taljard for kindly reviewing the Northern Sotho Rainbow Passage.

Statement of Ethics

This study protocol was reviewed and approved by the Research Ethics Committee of the Faculty of Humanities at the University of Pretoria (HUM018/0221). Written informed consent was obtained from each participant.

Declaration of Interest

The authors report there are no competing interests to declare.

Funding Sources

There are no funding sources to declare.

Author Contributions

AH conducted data collection and was the main contributor in writing the manuscript.

BP conceptualised the method, contributed to interpretation of results, and edited the manuscript.

MG conducted the statistical analysis and assisted with the presentation of results.

JvdL conceptualised the method, contributed to interpretation of results, and edited the manuscript.

Data Availability Statement

Data will be published and managed in an institutional Research Data Management system; an accredited open data repository.

References

1 Van der Merwe A, Le Roux M. Idiosyncratic sound systems of the South African Bantu languages: Research and clinical implications for speech-language pathologists and audiologists.

2 Ferrand CT. Voice disorders: Scope of theory and practice. Pearson Higher Ed; 2011

3 Godin KW, Hansen JH. Physical task stress and speaker variability in voice quality. EURASIP Journal on Audio, Speech, and Music Processing. 2015 Dec;2015(1):1-3.

4 Nevo L, Nevo C, Oliveira G. A comparison of vocal parameters in adult bilingual Hebrew-English speakers. InCoDAS 2015 Sep (Vol. 27, pp. 483-491). Sociedade Brasileira de Fonoaudiologia.

5 Lee B, Sidtis DV. The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. Speech, Language and Hearing. 2017 Jul 3;20(3):174-85.

6 Pépiot E, Arnold A. Cross-Gender Differences in English/French Bilingual Speakers: A Multiparametric Study. Perceptual and Motor Skills. 2021 Feb;128(1):153-77.

7 Pineda-Pérez E, Calvache C, Cantor-Cutiva LC. Bibliometric Analysis and Review of Literature on the Relationship Between Voice Production and Bilingualism. Journal of Voice. 2021 Aug 14.

8 Van Zyl M, Pillay B, Kritzinger A, Lekganyane M, Graham M. Significance of speech production errors on cross-linguistic processing in Sepedi-English individuals with bilingual aphasia: a case series analysis. Topics in Stroke Rehabilitation. 2019 May 19;26(4):294-306.

9 Lewis MP. Ethnologue: Languages of the world. SIL international; 2009.

10 Bokamba EG. Multilingualism and theories of second language acquisition in Africa. World Englishes. 2018 Sep;37(3):432-46.

11 Khokhlova I. Lingua Franca English of South Africa. Procedia-Social and Behavioral Sciences. 2015 Dec 5; 214:983-91.

12 Graham C. Fundamental frequency range in Japanese and English: The case of simultaneous bilinguals. Phonetica. 2014;71(4):271-95.

13 Ordin M, Mennen I. Cross-linguistic differences in bilinguals' fundamental frequency ranges. Journal of Speech, Language, and Hearing Research. 2017 Jun 10;60(6):1493-506.

14 Ng ML, Chen Y, Chan EY. Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers—a long-term average spectral analysis. Journal of Voice. 2012 Jul 1;26(4):e171-6.

15 Levi SV. Another bilingual advantage? Perception of talker-voice information. Bilingualism: Language and Cognition. 2018 May;21(3):523-36.

16 Hudson AI, Holbrook A. Fundamental frequency characteristics of young black adults: Spontaneous speaking and oral reading. Journal of Speech, Language, and Hearing Research. 1982 Mar;25(1):25-8.

17 Cantor-Cutiva LC, Bottalico P, Nudelman C, Webster J, Hunter EJ. Do voice acoustic parameters differ between bilingual english-spanish speakers and monolingual english speakers during english productions? Journal of Voice. 2021 Mar 1;35(2):194-202.

18 Cheng A. Cross-linguistic pitch differences in bilingual speakers. The Journal of the Acoustical Society of America. 2019 Mar;145(3):1825-.

19 Pinho CM, Jesus LM, Barney A. Weak voicing in fricative production. Journal of Phonetics. 2012 Sep 1;40(5):625-38.

20 Mennen I, Scobbie JM, de Leeuw E, Schaeffler S, Schaeffler F. Measuring language-specific phonetic settings. Second Language Research. 2010 Jan;26(1):13-41.

21 Coelho AC, Medved DM, Brasolotto AG. Hearing loss and the voice. An update on hearing loss. 2015 Dec 2:103-28.

22 Louw C, Swanepoel DW, Eikelboom RH, Myburgh HC. Smartphone-based hearing screening at primary health care clinics. Ear and Hearing. 2017 Mar 1;38(2):e93-100.

23 Fairbanks G. Voice and articulation drill book (Second). Harper & Brothers; 1960

24 Keating P, Kuo G. Comparison of speaking fundamental frequency in English and Mandarin. The Journal of the Acoustical Society of America. 2012 Aug;132(2):1050-60.

25 Goodglass H, Caplan E. The assessment of aphasia and related disorders. Lea & Febiger; 1972

26 Kreiman J, Sidtis D. Foundations of voice studies: An interdisciplinary approach to voice production and perception. John Wiley & Sons; 2011 Mar 21.

27 Theodore RM, Flanagan EG. Determinants of voice recognition in monolingual and bilingual listeners. Bilingualism: Language and Cognition. 2020 Jan;23(1):158-70.

28 Posel D, Hunter M, Rudwick S. Revisiting the prevalence of English: Language use outside the home in South Africa. Journal of Multilingual and Multicultural Development. 2020 Jun 9:1-3.

29 Department of Basic Education. CAPS for Intermediate Phase [Internet]. South Africa [updated 2011 January; cited 2022 August 22]. Available from: https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements(CAPS)/CAPSIntermediate.aspx

30 Pellegrino F, Coupé C, Marsico E. A cross-language perspective on speech information rate. Language. 2011 Sep 1:539-58.

31 Heugh K. Multilingual education policy in South Africa constrained by theoretical and historical disconnections. Annual Review of Applied Linguistics. 2013 Mar; 33:215-37.

32 Boersma P, Weenink D. Praat: Doing phonetics by computer (Version 5.4.560) [Computer program]. Available from: http://www.praat.org/

33 Yamauchi EJ, Imaizumi S, Maruyama H, Haji T. Perceptual evaluation of pathological voice quality: A comparative analysis between the RASATI and GRBASI scales. Logopedics Phoniatrics Vocology. 2010 Oct 1;35(3):121-8.

CHAPTER 4: Discussion

Chapter Aim: The study's results are summarised, and the ensuing theoretical and clinical implications are discussed. The strengths and weaknesses of the study are critically evaluated. Implications for future research precede the concluding statement.

4.1 Summary of Results

Acoustic vocal characteristics were contingent on task performance and differed across and within languages. Across languages, mean f_0 and intensity were significantly higher during the Northern Sotho picture description ($p = 0.002$) and reading passage ($p = 0.033$), respectively. Within English, reading elicited a significantly quicker speech rate than the picture description ($p = 0.003$) and monologue ($p = 0.028$). Within Northern Sotho, reading elicited a significantly higher mean f_0 than the monologue ($p = 0.028$). Perceptual voice quality, glottal attack, and resonance were comparable across English and Northern Sotho monologues. The discrepancy in significance across acoustic analysis and perceptual analysis of vocal characteristics highlights the value of both analyses for SLTs working with bilingual voice clients. Acoustic analysis may reveal vocal characteristics that are neither perceptually obvious to the SLT in spontaneous speech nor evident in only one of the bilingual's languages. The results corroborate with the growing body of research that suggests vocal characteristics manifest as a function of language and speech task in bilingual speakers (Lee et al., 2017; Pineda-Pérez et al., 2021).

4.1.1 Mean F_0

Mean f_0 did not manifest similarly across languages even when the same elicitation task was used ($p = 0.104$). There was no significant difference in mean f_0 between the

reading passage (p = 0.210) and monologue (p = 0.210) tasks in English and Northern Sotho. Describing a picture in Northern Sotho elicited significantly higher mean f_0 than describing a picture in English (p = 0.002). Mean f_0 was significantly higher during the Northern Sotho reading passage than during the Northern Sotho monologue (p = 0.028). Fluctuations in mean f_0 within the Northern Sotho samples corroborate with existing research that suggests tonal languages may be more susceptible to f_0 variations than non-tonal languages, like English (Cheng, 2020; Theodore et al., 2020).

4.1.2 Speech Rate

Speech rate was comparable across all speech tasks performed in Northern Sotho. Within English, reading elicited a significantly quicker speech rate than the picture description (p = 0.003) and monologue (p = 0.003). Speech rate differed significantly across English and Northern Sotho during the reading passage only (p = 0.002). The participants read significantly quicker in English than they did in Northern Sotho. Most participants reported speaking Northern Sotho at home with family and acquired English through preschool and primary school education. Acquiring English as an academic language necessitates reading regularly in English (Department of Basic Education, 2011). Reading aloud may underscore a language's acoustic properties as words are inadvertently hyperarticulated throughout the structured task (Cheng, 2020; Lee et al., 2017). The nature of reading aloud, paired with the participants' familiarity with reading aloud in English, may have contributed to different speech rates in the English reading passage. In other words, participants may have read quicker in English simply because they were more familiar with reading in English than their home language as English is the dominant language of learning and teaching in South

Africa. Insight into a bilingual speaker's language profile, in terms of acquisition and daily use, is valuable as it may inform the interpretation of differing vocal characteristics across speech tasks and languages (Pineda-Pérez et al., 2021). The degree of familiarity with task performance in specific languages may reveal differing acoustic measurements across and within a bilingual's languages, as shown in the current study.

4.1.3 Intensity

Intensity was similar within the English and Northern Sotho samples, corroborating with Lee et al.,' (2017) findings. Intensity differed significantly across English and Northern Sotho during the reading passage only. The intensity was significantly greater when reading in Northern Sotho than in English ($p = 0.033$). Intonational accents, like those in Northern Sotho, specify f_0 and intensity targets (Tilsen, 2016). This intrinsic feature of Northern Sotho was hypothesised to induce greater intensity and higher mean f_0 across all Northern Sotho speech tasks. However, only the reading passage ascribed to the anticipated trend. This reiterates that (1) the language effect on voice cannot be solely delineated based on the intrinsic features of languages and (2) assessing vocal characteristics across speech tasks can provide SLTs with a holistic overview of a bilingual client's voice profile.

4.2 Implications

4.2.1 Theoretical Implications

Across all domains of speech therapy, SLTs need a comprehensive understanding of the domain's typical development, associated influences, and presentation to encompass best-practice while working with individuals that present with the atypical

counterpart (Shipley et al., 2016). The physiology of monolingual voice production is relatively well-understood (Zhang, 2016). An understanding of the cross-linguistic effect on voice is emerging in bilingual speakers (Pineda-Pérez et al., 2021). A robust understanding of the language-effect on typical voice production is needed to understand the implications of bilingualism on disordered voice production. Current information about the language-effect on typical voice should be presented to undergraduate SLT students to align the curriculum with the latest research and promote best-practice (Department of Health, 2014). This is especially relevant in linguistically-diverse contexts, like South Africa, where SLTs are trained, and work with heterogenous populations.

4.2.3 Clinical Implications

It is essential to investigate the bilingual individual's language profile when rendering early communication, language, literacy, aural-rehabilitation, and neuro-rehabilitation services (Shipley et al., 2016). Insight into a bilingual speaker's language acquisition and daily use are also valuable when working with persons with voice disorders. Considerations that clinically relate to the study are presented, however, continued research is needed to either support or refute these hypotheses:

4.2.3.1 Assessment. SLTs could consider the following when assessing bilingual voice clients:

- ***Elicitation procedures.*** The current study's findings reiterate that vocal characteristics manifest as a function of a language- and task-effect in bilingual speakers. A voice assessment typically requires the client to perform a variety of speech tasks; irrespective of whether they are a monolingual or multilingual speaker (Ferrand, 2012). It may be viable for SLTs to assess bilingual voice

clients in both languages across speech tasks, as in a language or fluency assessment (Owens, 2014; van Borsel et al., 2011).

- ***Normative data.*** A normative database for vocal characteristics across languages other than English was not identified by the researchers. A bilingual voice client's acoustic measurements may deviate from the currently available vocal norms. Therefore, seemingly atypical acoustic measurements obtained during a voice assessment, that occurred in a language other than English, may simply be typical of that language, yet the specific language is not represented in the current voice norms. Current voice norms should be used with caution when working with bilingual voice clients (Altenberg et al., 2006). Continued research is needed to establish norms for vocal characteristics in languages other than English.

4.2.3.2 Treatment. SLTs could consider the following when treating bilingual voice clients:

- ***Counselling.*** SLTs may use the latest research to counsel bilingual voice clients about the language-effect on voice. Knowledge about the cross-linguistic effect on voice may help the client understand the rationale for their individualised therapy plan and promote PCC (Mophosho, 2018).
- ***Unique goals for bilingual voice clients.*** The parameters of the speaker's model voice, or ideal voice, may differ according to the spoken language. Identifying and transferring ideal vocal characteristics from one language to the other may be a worthwhile goal for bilingual voice clients and remain irrelevant for monolingual voice clients.

- ***Combining specific languages and speech tasks.*** A bilingual speaker's voice difficulties may be evident in specific combinations of languages and tasks, yet to a lesser degree in other combinations. Subsequently, combinations of specific languages and speech tasks may facilitate desired voice outcomes in bilingual speakers. For example, listening to oneself read with the ideal intensity in one language may increase the bilingual speaker's awareness and motivation to achieve the desired intensity in the other language. This consideration is consistent with an asset-based approach, whereby the SLT uses the client's existing strengths or skills to facilitate acquisition of a desired outcome (Shipley et al., 2016).
- ***Bilingual speakers' languages dynamically influence voice.*** SLTs and their bilingual voice clients may consider the bilingual individual's languages as dynamic influences on voice production. It is relatively effortless for proficient bilingual speakers to switch between their languages (Kroll et al., 2013). Switching languages, or code-switching, may cause the bilingual speaker's vocal characteristics to instantaneously change in a volitional manner. This may be empowering for the person with a voice disorder, as other extraneous influences, like anatomy, emotional state, and personality traits, cannot be instantaneously adapted.
- ***Code switching and the model voice.*** Contrasting vocal characteristics across differing speech tasks in each of the bilingual's languages may uniquely develop the speaker's intra-aural feedback loop needed to identify and establish their model voice in either language (Schenk et al., 2020). Therefore, treatment in two languages could afford bilingual voice clients more opportunities to identify their model voice, as it presents in specific language(s)

and tasks, in comparison with monolingual speakers. If vocal hyperfunction is more prominent in one language, sustained code-switching may promote hygienic voice production, buffer vocal hyperfunction, and prevent subsequent secondary organic pathologies for bilingual individuals with functional voice disorders. However, extensive research is needed to explore this notion.

4.3 Critical Evaluations

The strengths and limitations of the study were critically evaluated as discussed below.

4.3.1 Strengths

The method of the current study was based on the work of Lee et al., (2017) and Nevo et al. (2015). The stimuli used to elicit speech samples, and analyses software, were reliable and have been used in previous cross-linguistic studies (Keating et al., 2012; Lee et al., 2017; Nevo et al., 2015). Careful consideration was dedicated to minimising environmental errors, participant changes, and human error on behalf of the researcher throughout procedures. The validity of the data was preserved as the recordings were taken in the same sound-attenuated booth with the same equipment that was consistently organised. Participants were encouraged to take brief breaks to combat fatigue. The researcher cross-checked the acoustic measurements copied from Praat to Excel, as this was identified as the point wherein human error was most likely to occur. Integrating elements of established studies contributed to the current study's reliability and validity and yielded results comparable with existing research.

4.3.2 Limitations

The sample was purposefully limited to female speakers only to control for the confounding influence of gender on voice (Ferrand, 2012). Similarly, English-Northern Sotho was the only language pair investigated. Despite being viable regulations in the current study, restricting the sample to female English-Northern Sotho bilingual individuals reduced the generalizability of the results. Furthermore, the validity of the clinical implications presented is limited as the study was not conducted with bilingual speakers with voice disorders.

The researchers could not identify a standardized tool to evaluate linguistic proficiency in Northern Sotho for the current study. Subsequently, participants' self-reported information was used to determine proficiency in both languages. This may have limited homogeneity within the study's sample. Furthermore, all the participants acquired Northern Sotho before English. Including a second group of participants, who acquired English before Northern Sotho, could have allowed for more robust research (Pineda-Pérez et al., 2021). Individual differences, like age of acquisition and daily use, would be better-controlled. Comparisons could be made across the groups without order of acquisition of each language posing a confounding influence. In this manner, the language-effect on voice could be better-isolated.

The sample size (n = 19) was smaller than initially anticipated (n = 30). Restrictions imposed by the COVID-19 lockdown period reduced the number of prospective participants in Gauteng; the province where the study took place. Interested individuals contacted the researcher on social media after viewing the advertisement (Appendix C). However, many prospective participants relocated to different provinces

throughout lockdown, and subsequently could not participate in the study. The smaller sample size reduced the statistical power of the study (Leedy et al., 2019).

The original Rainbow Passage (Appendix E.1) is phonetically balanced to represent the variety of sounds and mouth movements used in unscripted English speech. The researchers cannot assert that the Northern Sotho reading passage (Appendix E.2) is phonetically balanced, despite being reviewed by a qualified linguist.

4.4 Future Research

Further insight into the cross-linguistic effect on voice is needed to support the development of standardized assessment tools, such as normative databases for different languages (Altenberg et al., 2006). Contextually-relevant assessment tools will augment PCC, and support SLTs to integrate culturally and linguistically appropriate services with evidence-based practice (Mophosho, 2018).

Larger cohort studies may be conducted to corroborate the results of this study. It is viable to conduct the study with language pairs of other official South African languages, such as isiZulu, isiXhosa, Sesotho, and Setswana. The order of acquisition of either language may be counter-balanced across groups of participants who are sequential bilinguals to allow for more robust research. F_0 variability may be investigated in future cross-linguistic studies that investigate Germanic-tonal language pairs. Tonal languages may be more susceptible to changes in f_0 variability (Cheng, 2020). F_0 variability comparisons may reveal language-based voice differences that would not present across other vocal parameters.

Pineda-Pérez et al.'s (2021) suggestion to include bilingual occupational voice users in cross-linguistic voice research is seconded. Voice is a primary work tool for many

professionals, including: singers, speakers, and teachers (Cataldo et al., 2020). Occupational voice users often perform their jobs in two or more languages (Sanssené et al., 2020). This additional vocal demand might influence the prevalence of voice disorders among this population. However, the language effect on voice has not yet been investigated in bilingual occupational voice users.

The cross-linguistic effect on voice should be investigated across male and female populations as the language and task effect on voice may manifest differently across genders (Pépiot et al., 2021). The interaction between language and gender may be relevant for SLTs working with transgender individuals. Current SLT practice in gender-affirming services primarily consists of voice masculinization or feminisation (Davies et al., 2015). The combination of specific languages and tasks may have the potential to facilitate desired voice outcomes in bilingual occupational voice users, transgender people, and individuals with voice disorders. However, extensive cross-linguistic voice research is required to explore this notion.

4.5 Conclusions

The multifaceted nature of voice production has been highlighted by this study. A language- and task- effect was evident in female English-Northern Sotho bilingual individuals. Acoustic vocal characteristics were contingent on the study's interaction between language and speech tasks across and within languages. English and Northern Sotho did not interact with the speech tasks in the same manner. The findings corroborate with studies that suggest language, task performance, and vocal characteristics should not be considered mutually exclusive entities (Pineda-Pérez et al., 2021).

The results of the study offer insight to benefit bilingual individuals receiving SLT services. Acoustic analysis may reveal significant vocal characteristics that are neither perceptually obvious to the SLT in spontaneous speech nor evident in only one of the bilingual's languages. An understanding of the bilingual individual's language profile, and the intrinsic features of each language, can inform the interpretation of differing vocal characteristics across speech tasks. SLTs could consider possible variations in vocal performance across speech tasks, and in different languages, to integrate linguistically appropriate services with PCC across populations.

REFERENCES

Altenberg, E. P., & Ferrand, C. T. (2006). Perception of individuals with voice disorders by monolingual English, bilingual Cantonese-English, and bilingual Russian-English women. *Journal of Speech, Language, and Hearing Research: Jslhr*, *49*(4), 879–887. https://doi.org/10.1044/1092-4388(2006/063)

Ansaldo, A. I., & Saidi, L. G. (2014). Aphasia therapy in the age of globalization: cross-linguistic therapy effects in bilingual aphasia. *Behavioural Neurology*, *2014*(1), 1–10. https://doi.org/10.1155/2014/603085

ASHA. (2015). Speech-Language Pathologists - Job Description and Career Information. Asha.org. https://www.asha.org/students/speech-language-pathologists/

Awan, S. N., & Mueller, P. B. (1996). Speaking fundamental frequency characteristics of white, African American, and Hispanic kindergartners. *Journal of Speech and Hearing Research*, *39*(3), 573–577. https://doi.org/10.1044/jshr.3903.573

Barcroft, J., & Sommers, M. S. (2014). Effects of variability in fundamental frequency on L2 vocabulary learning: A comparison between learners who do and do not speak a tone language. *Studies in Second Language Acquisition*, *36*(3), 423–449. https://doi.org/10.1017/s0272263113000582

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging*, *19*(2), 290–303. https://doi.org/10.1037/0882-7974.19.2.290

Boersma, P. & Weenink, D. (2019). *Praat: Doing phonetics by computer* (Version 5.4.56) [Computer program]. http://www.praat.org/

Bokamba, E. G. (2018). Multilingualism and theories of second language acquisition in Africa. *World Englishes*, 37(3), 432–446. https://doi.org/10.1111/weng.12330

Bradley, P. J. (2010). Voice Disorders: Classification. *Otorhinolaryngology, Head and Neck Surgery*, 555–562. https://doi.org/10.1007/978-3-540-68940-9_60

Cantor-Cutiva, L. C., Bottalico, P., Nudelman, C., Webster, J., & Hunter, E. J. (2021). Do voice acoustic parameters differ between bilingual English-Spanish speakers and monolingual English speakers during English productions? *Journal of Voice*, *35*(2), 194–202. https://doi.org/10.1016/j.jvoice.2019.08.009

Cataldo, E., Monteiro, L., & Soize, C. (2023). A Novel Source-Filter Stochastic Model for Voice Production. *Journal of Voice*, 37(1), 1–8. https://doi.org/10.1016/j.jvoice.2020.11.015

Cheng, A. (2020). Cross-linguistic f0 differences in bilingual speakers of English and Korean. *The Journal of the Acoustical Society of America*, *147*(2), 62 - 73. https://doi.org/10.1121/10.0000498

Coelho, A. C., Medved, D. M., & Brasolotto, A. G. (2015). Hearing Loss and the Voice. Update On Hearing Loss. https://doi.org/10.5772/61217

Davies, S., Papp, V. G., & Antoni, C. (2015). Voice and communication change for gender nonconforming individuals: Giving voice to the person inside. *International Journal of Transgenderism*, *16*(3), 117–159. https://doi.org/10.1080/15532739.2015.1075931

de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390. https://doi.org/10.3758/BRM.41.2.385

Department of Basic Education. (2011). *CAPS for Intermediate Phase* https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements(CAPS)/CAPSIntermediate.aspx

Department of Health. (2014). Health Professions Act: Regulations Relating to the Undergraduate Curricula and Professional Examinations in Speech-Language

Therapy.https://www.gov.za/sites/default/files/gcis_document/201501/38083gon782.pdf

Deumert, A. (2010). Tracking the demographics of (urban) language shift-an analysis of South African census data. *Journal of Multilingual and Multicultural Development*, *31*(1), 13–35. https://doi.org/10.1080/01434630903215125

DiLollo, A., & Favreau, C. (2010). Person-centered care and speech and language therapy. *Seminars in Speech and Language*, *31*(02), 090–097. https://doi.org/10.1055/s-0030-1252110

Fairbanks, G. (1960). *Voice and articulation drillbook* (Second). Harper & Brothers.

Ferrand, C. T. (2012). *Voice disorders: scope of theory and practice* (Ser.The Allyn & Bacon communication sciences and disorders series). Pearson.

Fouche, F., van der Merwe, A. (1999). Sepedi-toets vir spraakverstaanbaarheid. *South African Journal of Communication Disorders*, *46*(1), 25–35. https://doi.org/10.4102/sajcd.v46i1.725

Godin, K. W., & Hansen, J. H. L. (2015). Physical task stress and speaker variability in voice quality. *Eurasip Journal on Audio, Speech, and Music Processing*, *2015*(1), 1–13. https://doi.org/10.1186/s13636-015-0072-7

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Lea & Febiger.

Graham, C. (2014). Fundamental frequency range in Japanese and English: The case of simultaneous bilinguals. *Phonetica*, *71*(4), 271–95. https://doi.org/10.1159/000381627

Hazan, V., & Baker, R. (2010). Does reading clearly produce the same acoustic-phonetic modifications as spontaneous speech in a clear speaking style? In DiSS-LPSS Joint Workshop 2010.

Honikman, B. 1964: Articulatory settings. In Abercrombie, D., Fry, D.B., MacCarthy, P.A.D., Scott, N.C. and Trim J.L.M., editors, In honour of Daniel Jones. London: Longman, 73–84.

Hudson, A.I., Holbrook, A. 1982. Fundamental frequency characteristics of young black adults: spontaneous speaking and oral reading. *Journal of Speech and Hearing Research*, 25: 25–28. https://doi.org/10.1044/jshr.2501.25

Jackson, S. L. (2012). Research methods and statistics: a critical thinking approach. Wadsworth Cengage Learning.

Judith, F. K., & Ellen, B. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514. https://doi.org/10.1080/20445911.2013.799170

Kaur, P., Stoltzfus, J., & Yellapu, V. (2018). Descriptive statistics. *International Journal of Academic Medicine*, 4(1), 60–63. https://doi.org/10.4103/ijam.ijam_7_18

Keating, P., & Kuo, G. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *The Journal of the Acoustical Society of America*, *132*(2), 1050–60. https://doi.org/10.1121/1.4730893

Khokhlova, I. (2015). Lingua franca English of South Africa. *Procedia - Social and Behavioral Sciences*, *214*, 983–991. https://doi.org/10.1016/j.sbspro.2015.11.689

Khokhlova, I. N. (2007). Characteristics of South African Lexical Component in Modern English and Russian Languages. [Candidate dissertation]. Petropavlovsk-Kamchatsky: Vitus Bering Kamchatka State University Publishing House

Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: an interdisciplinary approach to voice production and perception*. Wiley-Blackwell.

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. Bilingualism: Language and Cognition, 13(3), 373–381. https://doi.org/10.1017/s136672891000009x

Lavan, N., Burton, A. M., Scott, S. K., & McGettigan, C. (2018). Flexible voices: identity perception from variable vocal signals. *Psychonomic Bulletin & Review*, *26*(1), 90–102. https://doi.org/10.3758/s13423-018-1497-7

Lee, B., & Sidtis, D. V. L. (2017). The bilingual voice: vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, *20*(3), 174–185. https://doi.org/10.1080/2050571X.2016.1273572

Leedy, P. D., & Ormrod, J. E. (2016). *Practical research: planning and design* (11th ed.). Pearson.

Levi, S. V. (2017). Another bilingual advantage? Perception of talker-voice information. *Bilingualism: Language and Cognition*, 21(3), 523–536. https://doi.org/10.1017/s1366728917000153

Lewis, M. P (Ed.). (2009). Ethnologue: Languages of the world (16th ed.) SIL International. Online version: Retrieved from https://www.ethnologue.com/16

Louw, C., Swanepoel, D. W., Eikelboom, R. H., & Myburgh, H. C. (2017). Smartphone-based hearing screening at primary health care clinics. *Ear and Hearing*, *38*(2), 100. https://doi.org/10.1097/AUD.0000000000000378

Malan, R. (2016). *Syllabic tone variation by Sepedi speakers with dysarthria* [Unpublished Master's dissertation] University of Pretoria, Pretoria, South Africa.

Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: a comparison of English and German. *The Journal of the Acoustical Society of America*, *131*(3), 2249–60. https://doi.org/10.1121/1.3681950

Mennen, I., Scobbie, J., de Leeuw, E., Schaeffler, S., & Schaeffler, F. (2010). Measuring language-specific phonetic settings. *Second Language Research*, *26*(1), 13–41. https://doi.org/10.1177/0267658309337617

Mophosho, M. (2018). Speech-language therapy consultation practices in multilingual and multicultural healthcare contexts: current training in South Africa. *African Journal of Health Professions Education*, *10*(3), 145–147. https://doi.org/10.7196/AJHPE.2018.v10i3.1045

Nasser, R., & Abolfazl, S. (2006). An introduction to speech sciences (acoustic analysis of speech). *Iranian Rehabilitation Journal*, *4*(4), 5–14.

Nevo, L., Nevo, C., & Oliveira, G. (2015). A comparison of vocal parameters in adult bilingual Hebrew-English speakers. *CoDAS*, *27*(5), 483–491. https://doi.org/10.1590/2317-1782/20152015096

Ng, M. L., Chen, Y., & Chan, E. Y. K. (2012). Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers—a long-term average spectral analysis. *Journal of Voice*, *26*(4), 171 - 176. https://doi.org/10.1016/j.jvoice.2011.07.013

Ordin, M., & Mennen, I. (2017). Cross-linguistic differences in bilinguals' fundamental frequency ranges. *Journal of Speech, Language, and Hearing Research: Jslhr*, *60*(6), 1493–1506. https://doi.org/10.1044/2016_JSLHR-S-16-0315

Owens, R. E. (2014). *Language disorders: a functional approach to assessment and intervention* (6th ed., Ser. Allyn & bacon communication sciences and disorders series). Pearson.

Pellegrino, F., Coupe, C., & Marsico, E. (2011). Across-language perspective on speech information rate. *Language*, *87*(3), 539–558. https://doi.org/10.1353/lan.2011.0057

Penn, C., & Armstrong, E. (2017). Intercultural aphasia: new models of understanding for indigenous populations. *Aphasiology*, *31*(5), 563–594. https://doi.org/10.1080/02687038.2016.1213788

Pépiot Erwan. (2015). Voice, speech, and gender. *Corela*, *(20150629)*. https://doi.org/10.4000/corela.3783

Pépiot, E., & Arnold, A. (2021). Cross-gender differences in English/French bilingual speakers: a multiparametric study. *Perceptual and Motor Skills*, *128*(1), 153–177. https://doi.org/10.1177/0031512520973514

Pineda-Pérez, E., Calvache, C., & Cantor-Cutiva, L. C. (2021). Bibliometric analysis and review of literature on the relationship between voice production and bilingualism. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2021.07.006

Pinho Cátia M.R, Jesus, L. M. T., & Barney, A. (2012). Weak voicing in fricative production. *Journal of Phonetics*, *40*(5), 625–638. https://doi.org/10.1016/j.wocn.2012.06.002

Posel, D., & Zeller, J. (2016). Language shift or increased bilingualism in South Africa: evidence from census data. *Journal of Multilingual and Multicultural Development*, *37*(4), 357–370. https://doi.org/10.1080/01434632.2015.1072206

Posel, D., Hunter, M., & Rudwick, S. (2020). Revisiting the prevalence of English: language use outside the home in South Africa. *Journal of Multilingual and Multicultural Development*, 43(8), 774–786. https://doi.org/10.1080/01434632.2020.1778707

Roberts, P. M. (2011). Chapter 15: Methodology Matters. Multilingual Aspects of Fluency Disorders, 353–370. https://doi.org/10.21832/9781847693570-017

Roberts, P., & Shenker, R. (2007). Assessment and treatment of stuttering in bilingual speakers. In E.G. Conture, & R. F. Curlee (Eds.), *Stuttering and related disorders of fluency* (pp. 183-209).Thieme Medical Publishers, Inc.

Rudwick, S. (2008). “Coconuts” and “oreos”: English-speaking Zulu people in a South African township. *World Englishes*, *27*(1), 101–116. https://doi.org/10.1111/j.1467-971X.2008.00538.x.

Sadat, J., Martin, C. D., Magnuson, J. S., Alario François-Xavier, & Costa, A. (2016). Breaking down the bilingual cost in speech production. *Cognitive Science*, *40*(8), 1911–1940. https://doi.org/10.1111/cogs.12315

Sanssené Claire, Bardi, J., & Welby-Gieusse, M. (2020). Prevalence and risk factors of voice disorders in French tour guides. *Journal of Voice*, *34*(6), 911–917. https://doi.org/10.1016/j.jvoice.2019.05.002

Schenck, A., Hilger, A. I., Levant, S., Kim, J. H., Lester-Smith, R. A., & Larson, C. (2020). The effect of pitch and loudness auditory feedback perturbations on vocal quality during sustained phonation. *Journal of Voice*, 37(1), 37–47. https://doi.org/10.1016/j.jvoice.2020.11.001

Schwartz, S. K. (2004). *The source for voice disorders: adolescent & adult*. LinguiSystems.

Shipley, K. G., & McAfee, J. G. (2016). *Assessment in speech-language pathology: a resource manual* (5th ed.). Cengage Learning.

Softtonic. (2019, July 1). Audacity (Version 2.3.2) [Computer program]. Audacity Team. http://www.audacityteam.org/

Statistics South Africa. (2012). Census 2011 Census in brief. http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf

Sunel, F., & Anita, van der M. (n.d.). Sepedi-toets vir spraakverstaanbaarheid. *South African Journal of Communication Disorders*, *46*(1), 25–35. https://doi.org/10.4102/sajcd.v46i1.725

Theodore, R. M., & Flanagan, E. G. (2020). Determinants of voice recognition in monolingual and bilingual listeners. *Bilingualism*, *23*(1), 158–170. https://doi.org/10.1017/S1366728919000075

Tilsen, S. (2016) A shared control parameter for F0 and intensity. *Proceedings of the International Conference on Speech Prosody 2016,* 1066-1070, doi: 10.21437/SpeechProsody.2016-219

Van der Merwe, A., & le Roux, M. (2014). Idiosyncratic sound systems of the South African Bantu languages: Research and clinical implications for speech-language pathologists and audiologists. *The South African Journal of Communication Disorders = Die Suid-Afrikaanse Tydskrif Vir Kommunikasieafwykings*, *61*(1), 8. https://doi.org/10.4102/sajcd.v61i1.86

van Zyl, M., Pillay, B., Kirzinger, A., Logandale, M., & Graham, M. (2019). Significance of speech production errors on cross-linguistic processing in Sepedi English individuals with bilingual aphasia: A case series analysis. *Topics in Stroke Rehabilitation*, *26*(4), 294–306. https://doi.org/10.1080/10749357.2019.1593612

Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to use what research design*. Guilford Press.

Wilsenach, C., & Makaure, P. (2018). Gender effects on phonological processing and reading development in Northern Sotho children learning to read in English: A case study of grade 3 learners. *South African Journal of Childhood Education*, *8*(1), 1–12. https://doi.org/10.4102/sajce.v8i1.546

Wong, P. C. M., Perrachione, T. K., Gunasekera, G., & Chandrasekaran, B. (2009). Communication disorders in speakers of tone languages: Etiological bases and clinical considerations. *Seminars in Speech and Language*, *30*(3), 162–173. https://doi.org/10.1055/s-0029-1225953

Yamauchi, E. J., Imaizumi, S., Maruyama, H., & Haji, T. (2010). Perceptual evaluation of pathological voice quality: a comparative analysis between the RASATI and GRBASI

scales. *Logopedics Phoniatrics Vocology*, *35*(3), 121–128. https://doi.org/10.3109/14015430903334269

Zhang, Z. (2016). Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *The Journal of the Acoustical Society of America*, *139*(4), 1493–1507. https://doi.org/10.1121/1.4944754

APPENDICES

Appendix A

Ethical Clearance

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Humanities
Fakulteit Geesteswetenskappe
Lefapha la Bomotho

Humanities 100
Since 1919

4 March 2021

Dear Miss A Hammann

Project Title:	Vocal characteristics across speech tasks in English-Northern Sotho bilingual individuals: A comparative study
Researcher:	Miss A Hammann
Supervisor(s):	Prof. Jeannie van der Linde Mrs SB Pillay
Department:	Speech Language Path and Aud
Reference number:	16032366 (HUM018/0221)
Degree:	Masters

I have pleasure in informing you that the above application was **approved** by the Research Ethics Committee on 25 February 2021. Data collection may therefore commence.

Please note that this approval is based on the assumption that the research will be carried out along the lines laid out in the proposal. Should the actual research depart significantly from the proposed research, it will be necessary to apply for a new research approval and ethical clearance.

We wish you success with the

project. Sincerely,

Prof Innocent Pikirayi

Deputy Dean: Postgraduate Studies and Research Ethics
Faculty of Humanities

UNIVERSITY OF PRETORIA

e-mail: PGHumanities@up.ac.za

Fakulteit Geesteswetenskappe
Lefapha la Bomotho

Research Ethics Committee Members: Prof I Pikirayi (Deputy Dean); Prof KL Harris; Mr A Bizos; Dr A-M de Beer; Dr A dos Santos; Ms KT Govinder Andrew; Dr P Gutura; Dr E Johnson; Prof D Maree; Mr A Mohamed; Dr I Noomé; Dr C Puttergill; Prof D Reyburn; Prof M Soer; Prof E Taljard; Prof V Thebe; Ms B Tsebe; Ms D Mokalapa

Appendix B

Informed Consent

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Humanities
Fakulteit Geesteswetenskappe
Lefapha la Bomotho

February 2021

Dear Participant

INFORMATION LEAFLET AND INFORMED CONSENT FOR PARTICIPANTS OF THE FOLLOWING STUDY:

Vocal characteristics across speech tasks in English-Northern Sotho bilingual individuals: A comparative study

1) INTRODUCTION

You are invited to volunteer for a research study conducted as part of a MA Speech-Language Pathology degree at the University of Pretoria. The information in this document is to help you to decide if you would like to participate. Before you agree to take part in this study, you should fully understand what is involved. If you have any questions, which are not fully explained in this document, do not hesitate to contact Amy Hammann at 076 441 6855 or amyhammann97@gmail.com. You should not agree to take part unless you are completely happy about all the procedures involved.

2) THE NATURE AND PURPOSE OF THIS STUDY

The aim of this study is to evaluate the effect of the language spoken and speech task performed on an individual's voice characteristics. By doing so, we wish to learn more about acquired factors which influence voice production.

3) EXPLANATION OF PROCEDURES AND WHAT WILL BE EXPECTED FROM PARTICIPANTS

If you want to participate, you have to be: female, between the ages of 18 and 65, and proficient in English and Northern Sotho. You will select a 1-hour timeslot for the data collection.

The data collection will take place in a quiet room or at the Voice Laboratory of the Department of Speech-Language Pathology and Audiology at the University of Pretoria. The researcher and each participant will adhere to guidelines outlined below to minimise the risk of spread of COVID-19 throughout data collection. The following precautions will be followed:

1. The researcher and all participants will complete a screening checklist prior to commencing with the data collection. Each person will indicate whether they have symptoms of COVID-19 and whether they have been in close contact with a person who tested positive for COVID-19 within the last 14 days. If a YES response is indicated, the session will be postponed.
2. All equipment and surfaces will be sanitized in between sessions. Alcohol-based sanitizer will be readily available throughout the sessions.
3. Social distance of 1-2 metres between people will be maintained throughout.
4. Masks will be worn where applicable. The participant will be required to remove, or lower, their mask while performing the speech tasks. This is necessary to preserve the integrity of the data as wearing a mask may impede voice production to an extent.
5. There will be 15-minute gaps between sessions to allow for ventilation and sanitizing between sessions with different participants.

On the day of the evaluation, a survey will be done to ensure that you have a healthy voice and a hearing screening will be conducted to confirm normal hearing. You will perform three different speech tasks. Each task should be done in English and in Northern Sotho, resulting in a total of six speech tasks being performed. The speech tasks will include: a reading passage, picture description and monologue. Your voice will be audio recorded while performing the speech tasks. The audio recordings will be used by the researcher to analyse your voice.

The following specific tests will be conducted for the research:

1. Each participant's hearing will be screened using the hearScreen screening application.

2. Each voice recording will be analysed using Praat software.

3. The GRBASI 4-point rating scale, and a nominal scale, will be used to evaluate the perceptual quality of each participant's voice.

4) RISKS AND DISCOMFORTS INVOLVED

No risks have been identified regarding participating in the study. Participants may take breaks between speech tasks to combat fatigue.

5) POSSIBLE BENEFITS OF THIS STUDY

Although you may not benefit directly, the study results may help us to better-understand the effect of language on voice production and the implications thereof in the field of speech-language therapy.

6) COMPENSATION

You will not be paid to take part in the study; no extra costs are expected to be concurred by you.

7) YOUR RIGHTS AS A RESEARCH PARTICIPAN

Your participation in this research study is entirely voluntary. You can withdraw from the study at any time without any consequences. Should you wish to withdraw, your data already collected will be excluded from the study. You have the right to access your data obtained for the study.

8) HAS THIS STUDY RECEIVED ETHICAL APPROVAL?

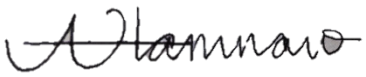
This study has received written approval from the Research Ethics Committee of the Faculty of Humanities at the University of Pretoria

9) CONFIDENTIALITY

All information obtained during the course of this study will be regarded as confidential. Each participant that is taking part will be provided with an alphanumeric coded number (Eg. P01). This will ensure that your information remains confidential. This code will only be known to the researcher and supervisors. Results will be published or presented in such a manner that you will remain unidentifiable. The hard copies of all your records will be kept in a locked facility at room 2-6 for 15 years at the Department of Speech-Language Therapy and Audiology, University of Pretoria. This data may be used for future research.

10) INFORMATION

If you have any questions about the study, please feel free to contact the researcher at 076 441 6855 or amyhammann97@gmail.com . Alternatively, you can contact the study supervisors, Prof. Jeannie van der Linde at jeannie.vanderlinde@up.ac.za or Mrs Bhavani Pillay at bhavani.pillay@up.ac.za.



Amy Hammann

Student

Prof. J. van der Linde

Supervisor

Mrs B. Pillay

Supervisor

CONSENT TO PARTICIPATE IN THE STUDY

I have read this information document and I understand the information discussed herein. I hereby agree to participate in the above-mentioned research project titled:

Vocal characteristics across speech tasks in English-Northern Sotho bilingual individuals: A comparative study. I have read the above information and understand what is required of me in this research study. I acknowledge that my results may be used confidentially for research purposes. I am aware that I participate voluntarily and that I may withdraw from the study at any time. I am aware that my results may be used for future research. I have received a signed copy of this informed consent agreement.

_____ _____

Participant name Date

_____ _____

Participant signature Date

_____ _____

Investigator's name Date

_____ _____

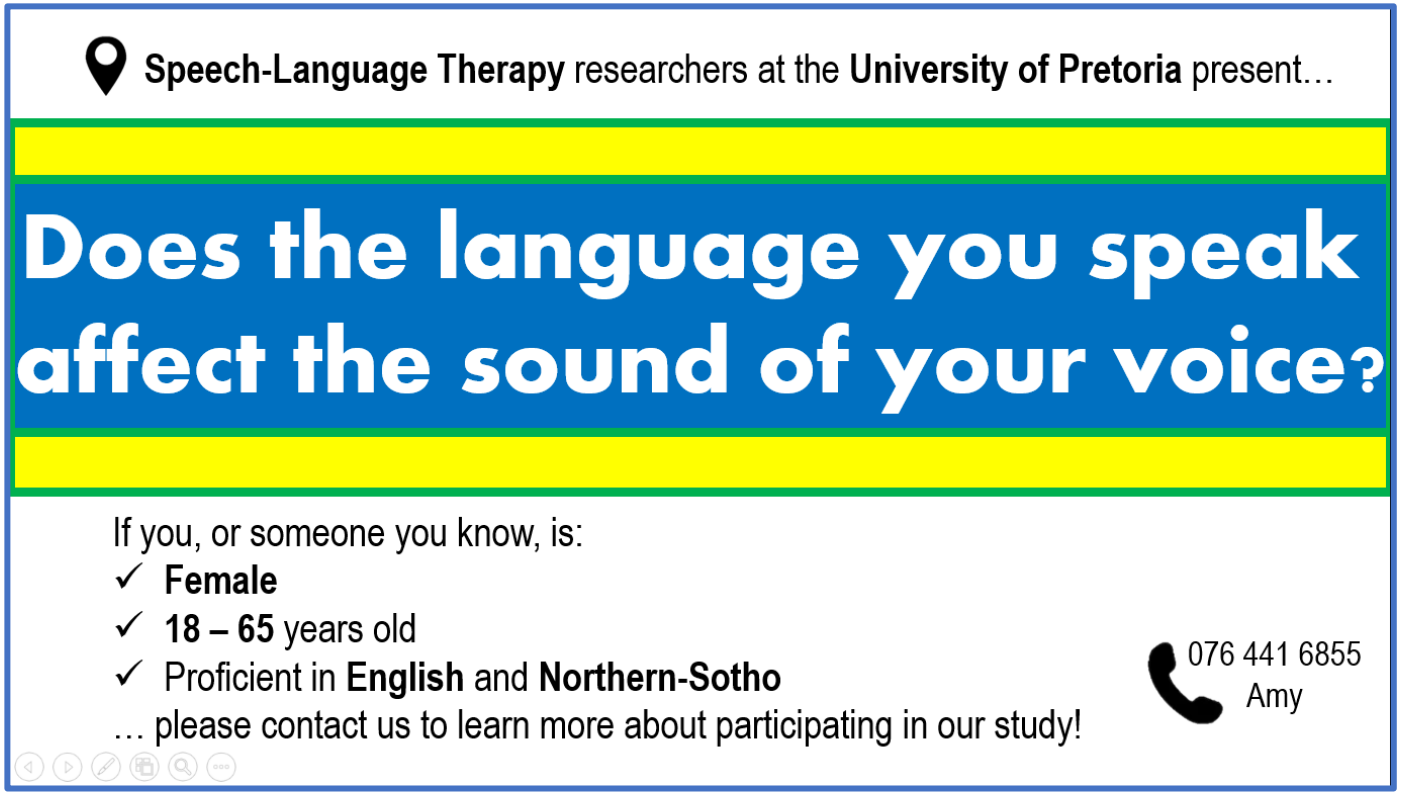
Investigator's signature Date

_____ _____

Witness name and signature Date

Appendix C

Advertisement



Appendix D

Participant Questionnaire

For Office Use

Questionnaire

Vocal characteristics across speech tasks in English-Northern- Sotho bilinguals: a comparative study

Please answer the questions by drawing a circle around the appropriate number in a shaded box or by writing your answer in the shaded space provided

Section 1: Participant Information

Participant code **(For office use only)** V1

Date of birth **(Please use dd/mm/yy)** V2

1. Do you feel that you have **normal hearing**?

Yes				1	V3
No				2	V4

2. Do you feel that you can **see well enough** (with spectacles if needed) to read?

Yes				1	V5
No				2	V6

3. Do you have any neurological condition which you feel may affect your voice?

Yes				1	V7
No				2	V8

Section 2: Voice Evaluation
(ASHA, 2018)

4. Have you ever been diagnosed with a **voice disorder**?

Yes	1	V9
No	2	V10

5. If yes, please describe: V11

6. Are you currently experiencing any symptoms of **flu**?

Yes	1	V12	
No	2	V13	

7. Have you ever suffered from **heartburn/reflux**?

Yes	1	V14	
No	2	V15	

8. If you answered yes in **question 4**, do you take any **medications** for heartburn/reflux?

Yes	1	V16	
No	2	V17	

9. Please specify the **type** and **dosage** of the medication if applicable.

	V18	

10. Are you currently suffering from any **allergies** that you feel may be influencing your voice production?

Yes	1	V19	
No	2	V20	

11.If you answered yes in **question 7**, are you currently taking any **medication** for allergies?

Yes	1	V21	
No	2	V22	

12. Please specify the **type** and **dosage** of the medication if **applicable.**

	V23	

13. Have you **ever** smoked cigarettes?

Yes	1	V24	
No	2	V25	

14. Do you **currently** smoke cigarettes?

Yes	1	V26	
No	2	V27	

15. If you answered yes to **question 10 or 11**, please describe the **duration** for which you have smoked. **(Please use y/mm)**

	V28	

16. Please indicate approximately **how many cigarettes** you smoke in an average day:

None	1	V29	

3	2	V30	
More than 3	3	V31	

17. Are you exposed to any of the following on a regular basis: (Indicate those applicable)

Smoke	1	V32	
Chemicals	2	V33	
Allergens	3	V34	
Severe temperature changes	4	V35	

18. For approximately **how long** are you exposed to the substance(s) indicated above?

Daily	1	V36	
Weekly	2	V37	
Monthly	3	V38	

19. Please indicate approximately **how much water** you drink in an average day:

None	1	V39	
500mls	2	V40	
1 litre or more	3	V41	

20. Do you feel that you have a **normal, healthy** voice?

Yes	1	V42	
No	2	V43	

Section 3: Language Preference
(Roberts et al., 2007; Lee et al., 2017)

21. What is your **home** language?
(Indicate those applicable)

Setswana	1	V45	
Northern-Sotho	2	V46	
Zulu	3	V47	
Shangaan	4	V48	
English	5	V49	
Afrikaans	6	V50	
Venda	7	V51	
Ndebele	8	V52	
Xhosa	9	V53	
Southern Sotho	10	V54	
SiSwati	11	V55	
Tsonga	12	V56	

Other (specify):

	V57	

22. What **other language(s)** do you speak?
(Indicate those applicable)

Setswana	1	V58	
Northern-Sotho	2	V59	
Zulu	3	V60	
Shangaan	4	V61	
English	5	V62	
Afrikaans	6	V63	
Venda	7	V64	
Ndebele	8	V65	
Xhosa	9	V66	
Southern Sotho	10	V67	
SiSwati	11	V68	
Tsonga	12	V69	

Other (specify):

	V70	

23. Do you consider yourself **bi/multilingual**?
i.e: A*ble to use different languages in writing, speaking and reading in different social contexts (*Roberts & Shanker, 2007)

Yes	1	V71	
No	2	V72	

24. In which language(s) do you **feel confident** communicating with other people who use that language as **their first language**? *(Indicate those applicable)*

Setswana	1	V73	
Northern-Sotho	2	V74	
Zulu	3	V75	
Shangaan	4	V76	
English	5	V77	
Afrikaans	6	V78	
Venda	7	V79	
Ndebele	8	V80	
Xhosa	9	V81	
Southern Sotho	10	V82	
SiSwati	11	V83	
Tsonga	12	V84	

Other (specify):

	V85	

25. When did you **first start** speaking English?

It is my first language	1	V86	
When I went to preschool	2	V87	
When I went to primary school	3	V88	
After primary school	4	V89	

26. In which **environment(s)** do you speak English in a typical day?

(Indicate those applicable)

Home	1	V90	
School	2	V91	
Work	3	V92	
Social	4	V93	
Other:			
		V94	

27. **With whom** do you speak English?
(Indicate those applicable)

Family	1	V95	
Coworkers	2	V96	
Friends	3	V97	
Other:			
		V98	

28. When did you **first start** speaking Northern-Sotho?

It is my first language	1	V99	
When I went to preschool	2	V100	
When I went to primary school	3	V101	
After primary school	4	V102	
Other:			
		V103	

29. In **which environment(s)** do you speak Northern-Sotho in a **typical** day?
(Indicate those applicable)

Home	1	V104	
School	2	V105	
Work	3	V106	
Social	4	V107	
Other:			
		V108	

30. **With whom** do you speak Northern-Sotho?
(Indicate those applicable)

Family	1	V109	
Coworkers	2	V110	
Friends	3	V111	
Other:			
		V112	

31. Please select the language(s) in which you are **most comfortable** communicating:

(Indicate those applicable)

Setswana	1	V113	
Northern-Sotho	2	V114	
Zulu	3	V115	
Shangaan	4	V116	
English	5	V117	
Afrikaans	6	V118	
Venda	7	V119	
Ndebele	8	V120	
Xhosa	9	V121	
Southern Sotho	10	V122	
SiSwati	11	V123	
Tsonga	12	V124	

Other (specify):

	V125	

32. Have you **personally noticed** any **changes** in your voice when speaking either English or Northern-Sotho?

Yes				1	V126	
No				2	V127	

33. If yes, please **describe further**?

	V128	

34. Is there any **additional information** that you feel would be relevant to inform the researchers about your **background information**, **voice** or **language** usage before participating in the study?

	V129	

Thank you for your time and co-operation!

Appendix E.1

Original Rainbow Passage

Instruction

Please familiarise yourself with this reading passage. When you feel comfortable, I will record you reading aloud. Please read in your most comfortable, natural way.

The Rainbow Passage

When the sunlight strikes raindrops in the air they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colours. These take the shape of a long round arch, with its path high above and its two ends apparently beyond the horizon. There is, according to a legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man is looking for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow. Throughout the centuries, men have explained the rainbow in various ways. Some have accepted it as a miracle without physical explanation. To the Hebrews, it was a token that there would be no more universal floods. The Greeks used to imagine that it was a sign from the gods to foretell war or heavy rain. The Norsemen considered the rainbow as a bridge over which the gods passed from earth to their home in the sky. Other men have tried to explain the phenomenon physically. Aristotle thought that the rainbow was caused by reflection of the sun's rays by the rain. Since then, physicists have found that it is not reflection, but refraction by the raindrops which causes the rainbow. Many complicated ideas about the rainbow have been formed. The difference in the rainbow depends considerably upon the size of the water drops, and the width of the coloured band increases as the size of the drops increases. The actual primary rainbow observed is said to be the effect of superposition of a number of bows. If the red of the second bow falls upon the green of the first, the result is to give a bow with an abnormally wide yellow band, since red and green lights when mixed form yellow. This is a very common type of bow, one showing mainly red and yellow, with little or no green or blue.

Appendix E.2

Northern Sotho Rainbow Passage

Instruction

Please familiarise yourself with this reading passage. When you feel comfortable, I will record you reading aloud. Please read in your most comfortable, natural way.

Tema ya molalatladi

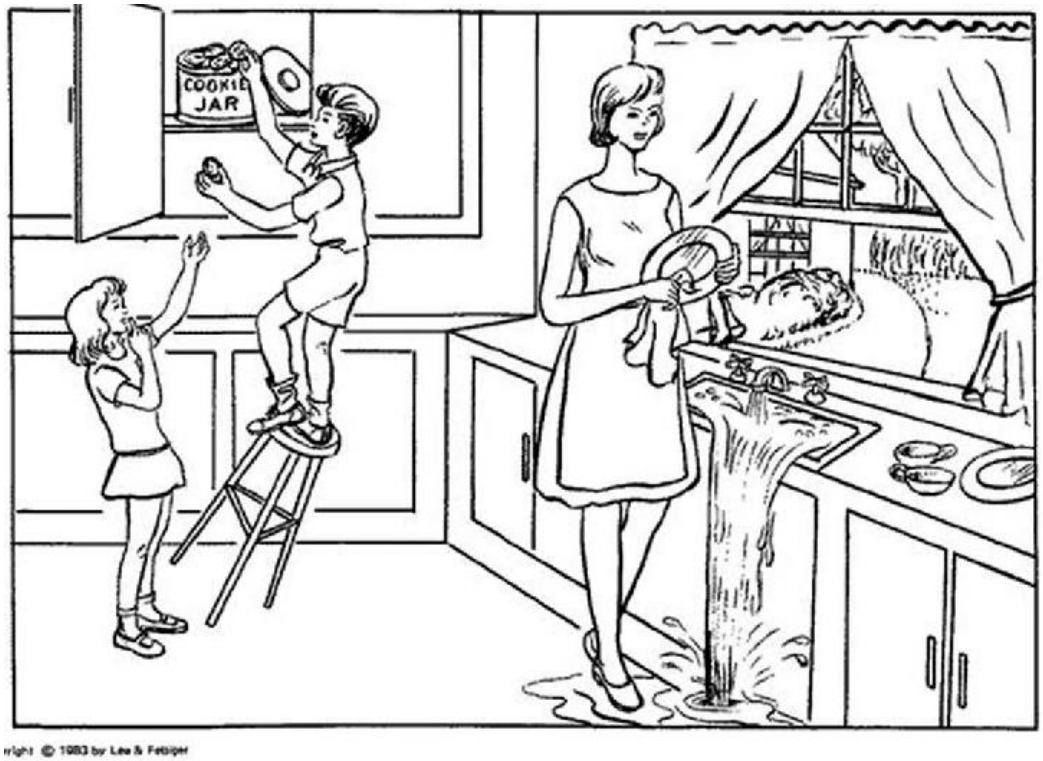
Ge masahledi a letšatši a tia marothedi a pula moyeng di latela molao wa go swana le wa phoresimi tša bopa molalatladi. Molalatladi ke karologanyo ya seetša se sešweu go mebala ye botse ye mentši. Se se tšea sebopego se setelele sa seripa sa nkgokolo, ka mohlala wo o lego godimodimo mme mafelelo a wona a bonala ka morago ga bogomapono. Go na le, go ya ka setsebalegi, pitša ye e belago ya gauta ka ntlheng ye nngwe. Batho ba nyaka, eupša ga go na yo a e hwetšago. Ge monna a nyaka selo se a ka se tsoge a se fihleletše, bagwera ba gagwe ba re o nyaka pitša ya gauta ka nthleng ya molalatladi. Ka gohle dingwagakgolong batho ba hlalošitše molalatladi ka ditsela tša go fapana. Ba bangwe ba e amogetše bjalo ka mohlolo ntle le thlalošo ya tlhago. Go Baheberu, e be e le sešupetšo sa gore go ka se hlwe go eba mafula mo gohle. Bagerika e be ba eleletša gore ke seka go tšwa go badimo go ba tsebišetšapele ka ntwa goba pula ye šoro. Banosemane ba tšea molalatladi bjalo ka leporogo leo badimo ba sepetšego ka lona go tloga lefaseng go ya legaeng la bona godimo. Ba bangwe banna ba lekile go hlaloša tiragalo ka sebopego. Aristotle o naganne gore molalatladi o hlolwa ke ponagatšo ya masahledi a letšatši ke pula. Go tloga fao, boramahlale ba hweditše gore ga se ponagatšo, eupša ke ponagatšophetogo ya marothedi a pula yeo e hlolago molalatladi. Dikgopolo tša go raragana tše ntši ka ga molalatladi di hlamilwe. Phapano ya molalatladi e laolwa kudu ke bogolo bja marothedi a meetse, gape bophara bja lepanta la mebala le gola ge bogolo bja marothedi bo oketšega. Molalatladi wa mathomo wa go lekolwa go thwe ke khuetšo ya peakanyo ye e kgethegilego ya mesebe ya palo ye e itšego. Ge bohubedu bja bora ya bobedi bo ka wela botala bja wa mathomo, poelo ke go fa bora ya lepanta la bokoto bja go se tlwaelege le le serolwana, ka gobane seetša se sehubedu le se setala ge di hlakantšwe di dira boserolwana. Se ke mohuta wa bora wo o tlwaelegilego, wo o laetšago fela bohubedu le boserolwana, ka bonnnyane goba go se tšweletše botala le mmala wa leratadima

Appendix E.3

The Cookie Theft Picture

Instruction A: *Take a few moments to look at the picture. Let me know when you are ready to describe it to me in* ***English****. Try to speak in your most comfortable, natural way. I will record you while you speak.*

Instruction B: *Take a few moments to look at the picture. Let me know when you are ready to describe it to me in* ***Northern Sotho****. Try to speak in your most comfortable, natural way. I will record you while you speak.*



Appendix E.4

Monologue Elicitation Topics

Topic 1 Option A:

Speech Task 1:

*I am going to record you while you speak for about 60 seconds. Tell me about something that you love about South Africa and why. Please speak in **English**.*

Speech Task 2:

*I am going to record you while you speak for about 60 seconds. Tell me about something that you love about South Africa and why. Please speak in **Northern Sotho**.*

Topic 1 Option B:

Speech Task 1:

*I am going to record you while you speak for about 60 seconds. Tell me about something that you love about South Africa and why. Please speak in **Northern Sotho**.*

Speech Task 2:

*I am going to record you while you speak for about 60 seconds. Tell me about something that you love about South Africa and why. Please speak in **English**.*

Topic 2 Option A:

Speech Task 1:

*I am going to record you while you speak for about 60 seconds. Tell me about your family. Please speak in **English**.*

Speech Task 2:

*I am going to record you while you speak for about 60 seconds. Tell me about your family. Please speak in **Northern Sotho**.*

Topic 2 Option B:

Speech Task 1:

*I am going to record you while you speak for about 60 seconds. Tell me about your family. Please speak in **Northern Sotho**.*

Speech Task 2:

*I am going to record you while you speak for about 60 seconds. Tell me about your family. Please speak in **English**.*

Appendix F

Perceptual Analysis Data Sheets

** Recordings are randomised. The recording number must **NOT** correspond with the participant code. E.g. : P01 ≠ E_1

RECORDING	G	R	B	A	S	I	Glottal Attack			Resonance		
E_1							h	s	a	n	t	o
E_2							h	s	a	n	t	o
E_3							h	s	a	n	t	o
E_4							h	s	a	n	t	o
E_5							h	s	a	n	t	o
E_6							h	s	a	n	t	o
E_7							h	s	a	n	t	o
E_8							h	s	a	n	t	o
E_9							h	s	a	n	t	o
E_10							h	s	a	n	t	o
E_11							h	s	a	n	t	o
E_12							h	s	a	n	t	o
E_13							h	s	a	n	t	o
E_14							h	s	a	n	t	o
E_15							h	s	a	n	t	o
E_16							h	s	a	n	t	o
E_17							h	s	a	n	t	o
E_18							h	s	a	n	t	o
E_19							h	s	a	n	t	o

GRBASI scale: 0 = normal; 1 = slight; 2 = moderate; 3 = severe (Yamauchi et al., 2010).

Nominal scale: h = harsh; s = soft; a = adequate; n = nasal; t = throaty; o = oral (Nevo et al., 2010)

RECORDING	G	R	B	A	S	I	Glottal Attack			Resonance		
NS_1							h	s	a	n	t	o
NS_2							h	s	a	n	t	o
NS_3							h	s	a	n	t	o
NS_4							h	s	a	n	t	o
NS_5							h	s	a	n	t	o
NS_6							h	s	a	n	t	o
NS_7							h	s	a	n	t	o
NS_8							h	s	a	n	t	o
NS_9							h	s	a	n	t	o
NS_10							h	s	a	n	t	o
NS_11							h	s	a	n	t	o
NS_12							h	s	a	n	t	o
NS_13							h	s	a	n	t	o
NS_14							h	s	a	n	t	o
NS_15							h	s	a	n	t	o
NS_16							h	s	a	n	t	o
NS_17							h	s	a	n	t	o
NS_18							h	s	a	n	t	o
NS_19							h	s	a	n	t	o

** Recordings are randomised. The recording number must **NOT** correspond with the participant code. E.g. : P01 ≠ NS_1

GRBASI scale: 0 = normal; 1 = slight; 2 = moderate; 3 = severe (Yamauchi et al., 2010)

Nominal scale: h = harsh; s = soft; a = adequate; n = nasal; t = throaty; o = oral (Nevo et al., 2015)

Appendix G

Journal Submission Confirmation

FPL-2022-11-10 Manuscript submission confirmation

Manuscript: FPL-2022-11-10 - Vocal characteristics across English-Northern Sotho bilingual speakers: a comparative study
Authors: Amy Hammann (Corresponding Author), Bhavani Pillay (Co-author), Marien Alet Graham (Co-author), Jeannie Van der Linde (Co-author)
Date submitted: 2022-11-27

Dear Miss Hammann

Thank you very much for submitting the above manuscript. Please refer to the manuscript number in all correspondence concerning the manuscript as listed above.

Please ensure that all co-authors confirm that you have the authority to act on their behalf by having them click the verification link on the submission confirmation email they have received.

The manuscript will now be forwarded to our Editors and reviewers and we shall inform you as soon as a decision has been made by the editorial board.

The progress of your manuscript can be followed in the progress report that can be accessed from your account overview.

Sincerely,
The Editorial Office