AUTHOR	Saloranta, Antti; Haapanen, Katja; Peltola, Kimmo U.; Tamminen, Henna; Uwu-khaeb, Lannie; Peltola, Maija S.
TITLE	Vowel qualities in monosyllabic words in Namibian English
YEAR	2024
DOI	https://doi.org/10.1558/jmbs.26377
VERSION	final draft
CITATION	Saloranta, Antti, Katja Haapanen, Kimmo U. Peltola, Henna Tamminen, Lannie Uwu-khaeb, and Maija S. Peltola. 2024. "Vowel Qualities in monosyllabic words in Namibian English". *Journal of Monolingual and Bilingual Speech* 6 (2): 213-24. https://doi.org/10.1558/jmbs.26377.

Vowel qualities in monosyllabic words in Namibian English

Antti Saloranta
University of Turku, Finland
Email:antti.saloranta@utu.fi (Corresponding author)

Katja Haapanen
University of Turku, Finland

Kimmo U. Peltola
University of Turku, Finland

Henna Tamminen
University of Turku, Finland

Lannie Uwu-khaeb
University of Turku in Windhoek, Namibia

Maija S. Peltola
University of Turku, Finland

DOI: 10.1558/jmbs.26377
Submitted June 1, 2023. Accepted October 5, 2023.

Abstract
This study investigated how Namibian English speakers produce vowel qualities in controlled speech. Participants were 14 proficient Namibian English speakers, who produced monosyllabic English CVC words, presented in orthographic form. The first and second formant values were extracted from the speakers' productions to determine the vowel qualities in each word. The formant values produced by the Namibian English speakers were then compared to those produced by nine British English speakers, obtained from a previous study, in order to examine how Namibian English vowel qualities relate to British English vowel categories. The results showed a great deal of variation and overlap in the Namibian vowel categories, particularly in the close front vowels and open and close back vowels. This resulted in the formation of five to six clusters of vowel qualities. This is likely an effect of the varied linguistic backgrounds of the participants. Furthermore, the Namibian English vowels differed significantly from most British English vowels, especially in the *F2* formant. Some commonalities with earlier research into Namibian English vowel qualities was found in the fronting, backing and mergers of some vowel categories.

Keywords: speech production; vowel quality; Namibian English; world Englishes; English vowels

1. Introduction

Namibia, situated in southwestern Africa, gained its independence from South Africa in 1990. Since then, English has been the country's official language. It was chosen over the commonly spoken Afrikaans or any of Namibia's indigenous languages, as the former was considered to be a symbol of oppression, and choosing any of the latter would have gone against the idea of Namibian unity. Furthermore, English was thought to allow for wider communication with other African countries and the rest of the world (Norro, 2022).

Recently, it has been suggested that an independent variety of Namibian English, separate from South African English varieties, is emerging among World Englishes (Buschfeld & Kautzsch, 2014; Schröder et al., 2021; Stell, 2022). Some phonetic features separate Namibian English from other English varieties, and more specifically from South African English (Buschfeld & Kautzsch, 2014; Schröder et al., 2021). These include the TRAP-DRESS merger, where the typically separate /æ/ and /e/ vowel categories are realized with similar qualities, and the WORK-NURSE split, where two separate vowel qualities are used instead of the /ɜ/ vowel. However, some suggest that the production of Namibian English vowels varies between speakers and is ethnically conditioned, meaning that there might be multiple local sub-varieties of Namibian English (Schröder et al., 2021). Others propose that a single, more neutral, high-status local variety of Namibian English might be forming among speakers of different local languages, uniting the ethnically distinct Namibian English varieties (Stell, 2022). However, rather little research on the phonetic features of Namibian English is available, with literature on English in Africa often focusing on South African varieties, or making general statements about English spoken by native speakers of Bantu languages (e.g., Nelson, 2020).

In addition to English, at least 28 different indigenous languages, as well as some Germanic languages, are spoken across Namibia, making it a linguistically diverse country. According to the 2011 census, the most common indigenous languages among the population of 2.1 million are Oshiwambo (49% of households), Khoekhoegowab (11%), Kavango (9%) and Otjiherero (9%; see Norro, 2022). Though English is the official language, most Namibians are multilingual, with one or many local language(s) as their first language (L1) or home language (Norro, 2022). Ten of the local languages have official status as school-languages, and are used in classrooms in the first years of school, but the medium of instruction changes to English in the fourth grade at the latest (Norro, 2021; 2024). Most of the local languages belong to the Bantu and Khoisan families, which are typically described as having five (/i, e, a, o, u/) or seven (/i, e, ɛ, a, ɔ, o, u/) vowel categories in their phonological systems (Cruttenden, 1992; Kilarski & Dziubalska-Kołaczyk, 2012; Maddieson & Sands, 2019; Odden, 2015), whereas, for example, British English (received pronunciation, RP) has a much larger vowel inventory, with 12 monophthong vowels: /iː/, /ɪ/, /e/, /æ/, /ʊ/, /uː/, /ɔː/, /ʌ /, /ɒ/, /ɑː/, /ɜ/ and /ə/ (Deterding, 1997; Roach, 2004). As English is often learned as a second or third language, and Namibian English speakers are mostly multilingual, the phonological systems of Bantu and Khoisan languages could be expected to influence the phonetic features of Namibian English. This expectation is based on previous studies that have found Bantu influences on spoken Black South African English (see, e,g., Louw & de Wet, 2007; Makalela, 2013; Nelson, 2020).

This experiment investigated the quality of Namibian English (NamE) vowels in monosyllabic words produced by speakers of varying Bantu and Khoisan languages. Previous studies into Namibian English have largely focused on vowel tokens extracted from somewhat uncontrolled speech, such as interviews (Schröder et al., 2021) or a reading task of continuous text (Stell, 2022). The downside of this is that vowel tokens are often short and affected by coarticulation effects. Our use of individual words in controlled and similar acoustic contexts in this study allowed us to examine the qualities of the vowels in their steady-state forms, providing new information about the Namibian English vowel system. The study aimed to answer the following research questions. First, what are the acoustic vowel qualities produced by Namibian English speakers in controlled, monosyllabic English words? Second, how do the vowel qualities produced by Namibian English speakers relate to the British English (BrE) vowel qualities in the same words? The hypothesis, based on previous research on the issue (Buschfeld & Kautzsch, 2014; Schröder et al., 2021), was that the NamE vowel qualities might be influenced by the small vowel systems of the local Bantu

and Khoisan languages, likely resulting in overlapping productions of different vowel categories, such as the TRAP-DRESS merger or STRUT fronting, as discussed by Stell (2022). We further hypothesized that the vowel qualities would therefore be different from those of BrE. The purpose of this comparison was to have a well-established larger variant of English acting as an anchor point to which the NamE vowel qualities can be contrasted as its own system. As the overall number of speakers in the study is fairly low, this study focused on providing an overview of the vowel system of Namibian English, rather than making any predictions based on specific languages.

2. Materials and methods

2.1 Participants

The experiment included 14 participants (aged 21–25 years, mean age 22.1, nine females). All of them were either students or employees of the University of Namibia in Windhoek. They had learnt English when they were 3–7 years old and spoke it on a daily basis. The participants' self-reported first or strongest languages were Khoekhoegowab (*n* = 4), Oshiwambo (2), Otjiherero (3), Oshikwanyama (1), Rukwangali (1), Mbalangwe (1) and English (2). Participants were also asked to assess their skill level in the languages they spoke on a scale from 1 = *basics* to 5 = *native-level*. The mean self-assessed English skill level was 4.07. All participants reported speaking one to four languages in addition to their strongest language. These included Afrikaans (9), Portuguese (1), Spanish (1) and Silozi (1).

2.2. Stimuli and procedure

The stimuli were 20 monosyllabic English words (Table 1). The final consonant was voiceless in 10 words and voiced in 10, and the words formed minimal or subminimal pairs based on this difference (i.e., HEAT-HEED, HIT-HID, etc.). The same English words have been used in two previous studies as orthographic stimuli for L1 and L2 British English-speaking adults (Peltola et al., 2014) and as auditory stimuli for L2 British English-speaking children (Immonen et al., 2018). In this report, the vowels will be referred to by the stimulus name (i.e., 'the HEAT vowel'). More information on the stimuli can be found in Peltola et al. (2014).

Table 1. The 20 stimulus words used in the experiment. The final consonant was voiceless (10 words) or voiced (10 words). Each column represents a minimal or subminimal pair, with a voicing contrast in the final consonant.

Voiceless	heat	hit	bet	hat	foot	hoot	bought	hut	tot	heart
Voiced	heed	hid	bed	had	hood	who'd	board	hud	Todd	hard

The stimuli were displayed visually in orthographic form on a Dell Latitude 5320 laptop computer using a PowerPoint Presentation. During the experiment, each of the 20 words was shown three times, for a total of 60 words. After the participant started the experiment, the stimuli were shown on the screen automatically with a 3-second inter-stimulus interval (ISI). The order of the stimuli was semi-randomized to ensure that no word appeared twice in a row during the test. The experiment took about 5 minutes, with one self-paced break in the middle. The participants were asked to say each word as soon as it appeared on the screen, and their productions were recorded using Sanako Study Recorder software and a Beyerdynamic MMX300 headset microphone connected to a Deltaco UAC–03 sound card.

2.3. Analysis

Acoustic analysis of the participants' productions was performed using Praat (Version 6.2.20; Boersma & Weenink, 2022). To determine the quality of the produced vowels, the first (*F1*) and second (*F2*) formant values were extracted from the midpoint of the vowels using the Linear Predictive Coding (LPC) Burg algorithm. The maximum frequency was set individually at 5000–6000 Hz for each speaker. Individual average *F1* and *F2* values were then calculated for each word from the three repetitions produced during the experiment. For reference, the formant values were then compared to values extracted from the same words produced by nine L1 BrE speakers in an earlier study (for detailed information about the speakers, see Peltola et al., 2014), to examine whether the vowel qualities of NamE speakers were different from BrE speakers. An overall examination of the vowel system was then performed, outlining the main features of the NamE vowel system and any major differences between it and the BrE vowel categories. The formant values were then subjected to one-way ANOVAs using the SPSS Statistics (Version 27.0.1.0) software, to investigate whether the NamE vowels' *F1* and *F2* values were significantly different to the BrE vowels.

3. Results

Examination of the average formant values of the Namibian participants' productions (Figure 1) showed that the front vowels in HEAT-HEED, HIT-HID, BET-BED and HAT-HAD had been produced with virtually identical qualities in the voiced and the voiceless contexts. Slightly more differentiation could be seen in the back vowels, particularly in FOOT-HOOD and HOOT-WHO'D, with both showing higher *F2*s and higher *F1*s in the voiceless contexts, meaning that acoustically, the vowels were more central. There was also clear overlap between the two vowel categories, as the productions of HOOD and HOOT had essentially identical vowel qualities.

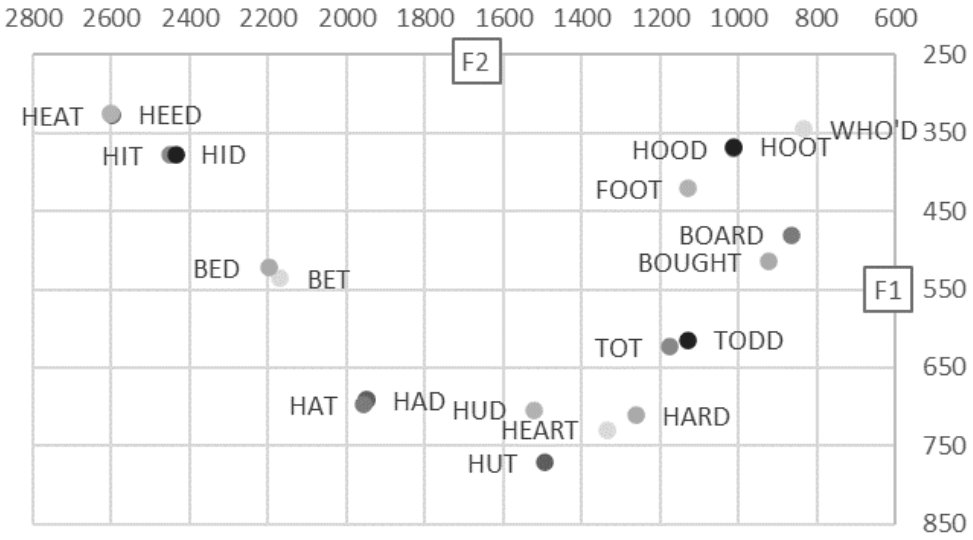
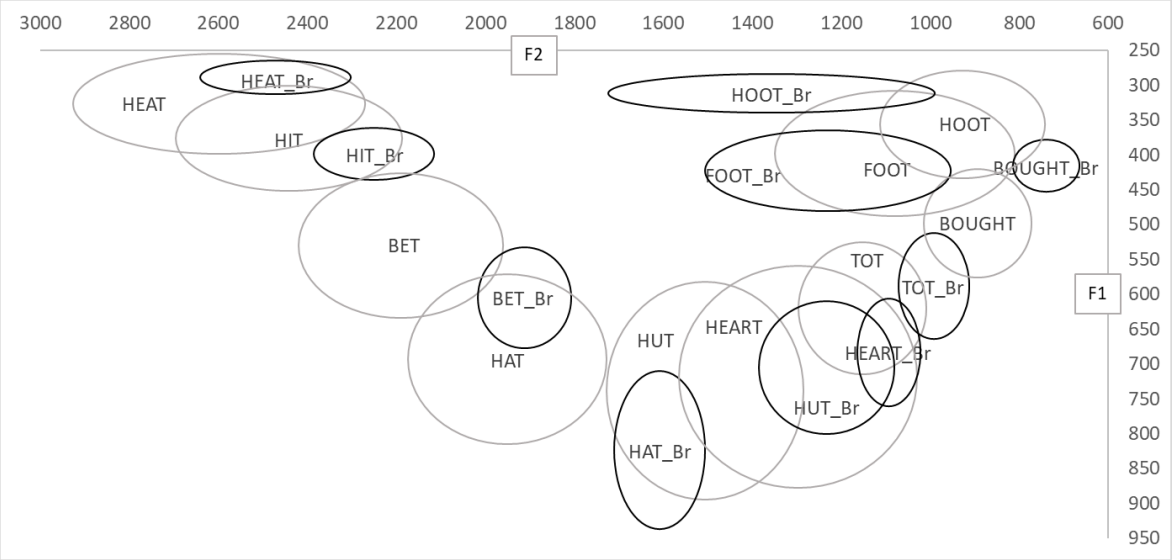


Figure 1. Mean formant values (Hz) for each stimulus word produced by the NamE speakers.

In order to get a better understanding of the variation and overlap in the vowels produced by the Namibian speakers, and to compare them to BrE productions of the same words, the mean productions of the voiceless and voiced contexts were combined in a vowel chart (Figure 2) For example, the mean *F1* and *F2* values in the chart represent the mean *F1* and *F2* of both HEAT and HEED together. This was done in order to reduce focus on voicing-based variation and to concentrate on overall variation in vowel quality. In Figure 2,

the word pairs are represented with the voiceless member of the pair (i.e., HEAT for HEAT-HEED. In addition, ellipses are used to depict the standard deviation of the formants. The mean values are at the center of the ellipses, and the radii are equivalent to one standard deviation of each formant, with the vertical radius representing *F1* and the horizontal *F2*.



[FIGURE 2 NEAR HERE]

Figure 2. Variation of vowel production by the NamE (green) and BrE (red) speakers with voiced and voiceless contexts combined. The vertical and horizontal radii of the ellipses represent one standard deviation of the *F1* and *F2* formants (Hz), respectively, with the mean at the center.

The chart showed that for the NamE speakers, there was a great deal of variation in the production of most of the vowels, particularly the front and open ones. In addition, several of the productions overlap notably, suggesting that the vowels in these words have been produced with somewhat similar formant values. Clusters of overlapping vowel areas can be seen in the close front vowels, open back vowels and close back vowels. The mid front BET-BED vowel and the open front HAT-HAD vowels are quite clearly separated, as are the HAT-HAD vowel and the open mid-back cluster (i.e., HUT-HUD, HEART-HARD and TOT-TODD). The BrE categories, on the other hand, were mostly consistent and clearly separated, with the exception of HUT-HUD and HEART-HARD, which overlap slightly, and the very small overlap between HEART-HARD and TOT-TODD.

Compared to the BrE speakers, the front and open vowels produced by the NamE speakers generally had more fronted *F2* values, and in the close back vowels the *F2* difference varied in both directions with regard to the BrE productions. *F1* variation was less consistent, with some vowels showing virtually the same *F1*s as in the BrE productions, and others being notably higher or lower.

In order to statistically compare the differences between the Namibian and the British speakers' productions, one-way ANOVAs (Group 2) were run between the two speaker groups for each formant in all of the stimulus words. Statistically significant differences, along with mean formant values for each word, can be seen in Table 2.

Table 2. Mean formant values and between-group one-way ANOVAs for all productions of the stimulus words for NamE and BrE speakers. Statistically significant differences between BrE and NamE productions are marked with an asterisk next to the NamE formant values.

	NamE speakers		BrE speakers		One-way ANOVAs	
	F1	***F2***	***F1***	***F2***	***F1***	***F2***
HEAT	328	2599	300	2452		
HEED	326	2597	287	2483		
HIT	378	2451*	420	2227		$F(1, 21) = 5.152$; $p = .034$; $\eta^2 = .197$
HID	379	2433	379	2269		
BET	536*	2167*	632	1898	$F(1, 21) = 5.799$; $p = .025$; $\eta^2 = .216$	$F(1, 21) = 10.456$; $p = .004$; $\eta^2 = .332$
BED	523	2196*	578	1934		$F(1, 21) = 8.664$; $p = .008$; $\eta^2 = .292$
HAT	698*	1955*	836	1602	$F(1, 21) = 8.025$; $p = .01$; $\eta^2 = .276$	$F(1, 21) = 17.111$; $p < .001$; $\eta^2 = .449$
HAD	692*	1946*	812	1612	$F(1, 21) = 4.732$; $p = .041$; $\eta^2 = .184$	$F(1, 21) = 18.172$; $p < .001$; $\eta^2 = .464$
FOOT	423	1132	461	1170		
HOOD	373	1031*	388	1294		$F(1, 21) = 4.71$; $p = .042$; $\eta^2 = .183$
HOOT	370	1026*	316	1384		$F(1, 21) = 7.852$; $p = .011$; $\eta^2 = .272$
WHO'D	346	838*	310	1329		$F(1, 21) = 22.495$; $p < .001$; $\eta^2 = .517$
BOUGHT	517*	925*	424	752	$F(1, 21) = 9.477$; $p = .006$; $\eta^2 = .311$	$F(1, 21) = 11.878$; $p = .002$; $\eta^2 = .361$
BOARD	483*	870*	412	729	$F(1, 21) = 7.67$; $p = .011$; $\eta^2 = .268$	$F(1, 21) = 11.292$; $p = .003$; $\eta^2 = .350$
HUT	772	1497*	738	1224		$F(1, 21) = 9.849$; $p = .005$; $\eta^2 = .319$
HUD	704	1522*	679	1242		$F(1, 21) = 10.357$; $p = .004$; $\eta^2 = .330$
TOT	624	1177*	610	1018		$F(1, 21) = 7.599$; $p = .012$; $\eta^2 = .266$
TODD	618	1132*	569	971		$F(1, 21) = 11.319$; $p = .003$; $\eta^2 = .350$
HEART	726	1337*	700	1103		$F(1, 21) = 5.688$; $p = .027$; $\eta^2 = .213$
HARD	711	1261	666	1085		

The results from the one-way ANOVAs showed that there was a statistically significant difference in at least one formant between the two speaker groups in all but three words: HID, FOOT and HARD. All the significant differences between NamE and BrE vowels in the tested words were found in at least the *F2*, and additionally in the *F1* for BET, HAT, HAD, BOUGHT and BOARD. The statistical findings therefore largely confirmed the observations from Figure 2, where *F2* differences were the more consistent factor separating the NamE and BrE productions.

4. Discussion

The purpose of this study was to examine the production of Namibian English vowels in controlled speech by speakers of various Namibian Bantu and Khoisan languages. The examination was performed using acoustic analysis of the vowel formants, and the results were statistically compared to productions of the same stimulus words by British English speakers.

Overall, the NamE vowels show a great deal of variation in both the *F1* and *F2* formants. This was somewhat expected, due to both the various linguistic backgrounds of the speakers, and the fact that the fundamental frequencies of the productions were not normalized in any way in this study.

As for the actual vowel qualities produced by the NamE speakers, only two vowels could be said to exist in a part of the vowel space largely unoccupied by any other category: BET-BED and HAT-HAD. All others showed a varying degree of overlap with one or more different categories, and it seemed that there were five main quality clusters in the productions, as shown by Figure 2: close front vowels, open back vowels and close back vowels, in addition to BET-BED and HAT-HAD. However, it could also be argued that the NamE BOUGHT-BOARD was its own category due to its tight variation and relatively small overlap with its neighboring categories. This would bring the number of different qualities to five or six, which is a typical number for several of the Bantu and Khoisan languages (Kilarski & Dziubalska-Kołaczyk, 2012; Maddieson & Sands, 2019; Odden, 2015). It is likely, however, that closer examination of the clusters based on language backgrounds would reveal varying differentiation patterns and reduce the variation in the categories, but this is outside the scope of this study. More speakers for each language would be needed to draw reliable language-specific conclusions. A noteworthy detail in Figure 1 was the complete overlap of the mean formant values for the different vowel categories in HOOT and HOOD. FOOT and WHO'D, the other members of the pairs, were clearly distinct from one another. This may be an effect of the test protocol itself: as the productions were elicited by reading the words aloud, the very similar orthographic forms of the words may have been a source of confusion.

Comparing the productions of the NamE and the BrE speakers showed that the formant values differed significantly between the two groups in the majority of the vowels, particularly the *F2*. The NamE speakers' *F2* showed more fronted values for the front and open vowels than the BrE speakers, and, conversely, more back values in the close back vowels, with the exception of the BrE BOUGHT-BOARD. High *F2* values could partially be explained by the slight bias towards female speakers in the NamE data, as higher fundamental frequency corresponds with higher formant values. However, in the close back vowels, the NamE speakers produced lower mean *F2* values than the BrE speakers, particularly for HOOT-WHO'D. Examination of earlier literature on Namibian English vowels does indeed show that this may be an effect of language background (Stell, 2022). Stell describes the realizations of English vowels produced by Namibian non-Afrikaners, and some shared findings in the current data include a weak FLEECE-KIT distinction (HEAT-HEED vs. HIT-HID in the current study), a closer DRESS-TRAP distinction (BET-BED vs. HAT-HAD), a fronted STRUT (HUT-HUD), and a backer FOOT-GOOSE vowel (FOOT-HOOD and HOOT-WHO'D). Again, a closer examination of specific language groups and even individual speakers, such as in Schröder et al. (2021), might reveal a more complicated pattern of similar and dissimilar vowel qualities, but this falls outside the scope of this study.

5. Conclusions

As hypothesized, the results of this study suggested that the English spoken by Namibians reflects, at least partially, the various linguistic backgrounds of the speakers. This was seen in

the significant overlap and large variation in several of the NamE categories. Significant differences between the vowels produced by Namibian English and British English speakers were also found, both in the actual formant values and the overall patterns the vowel categories form. Both of these findings are likely reflections of the relatively small number of vowel categories in the speakers' native languages, particularly in comparison to British English. The results also agreed with earlier research into Namibian English vowels. As the speaker numbers for individual languages in the current study were still rather small, no conclusions regarding any specific language could be reached. Future studies with larger speaker groups are needed to examine the effects of the different Bantu and Khoisan languages separately, and to find out what, if any, features of the vowel productions may be shared across linguistic boundaries, marking a specifically Namibian variant(s) of English.

Acknowledgments
The authors wish to thank the Kone Foundation for funding this work, Sanako Corp. for gifting the software used in the speech recordings, Professors Erkki Sutinen and Helvi Haikokolla for their invaluable help in Namibia, as well as all the speakers who participated in this study.

References

Boersma, P., & Weenink, D. (2022). *Praat* (Version 6.2.20). [Computer software]. http://www.fon.hum.uva.nl/praat/

Buschfeld, S., & Kautzsch, A. (2014). English in Namibia: A first approach. *English World-Wide: A Journal of Varieties of English*, 35(2), 121–160. https://doi.org/10.1075/eww.35.2.01bus

Cruttenden, A. (1992). Clicks and syllables in the phonology of Dama. *Lingua*, 86(2), 101–117. https://doi.org/10.1016/0024-3841(92)90031-D

Deterding, D. (1997). The formants of monophthong vowels in standard southern British English pronunciation. *Journal of the International Phonetic Association*, *27*(1–2), 47–55. https://doi.org/10.1017/S0025100300005417

Immonen, K., & Peltola, M. S. (2018). Finnish Children Producing English Vowels—Studying in an English Immersion Class Affects Vowel Production. *Journal of Language Teaching and Research*, 9(1), 27–33.

Kilarski, M., & Dziubalska-Kołaczyk, K. (2012). On extremes in linguistic complexity: Phonetic accounts of Iroquoian, Polynesian and Khoesan. *Historiographia Linguistica: International Journal for the History of the Language Sciences/Revue Internationale Pour l'Histoire Des Sciences Du Langage/Internationale Zeitschrift Für Die Geschichte Der Sprachwissenschaften*, *39*(2–3), 279–303. https://doi.org/10.1075/hl.39.2-3.05kil

Louw, P., & de Wet, F. (2007). The perception and identification of accent in spoken Black South African English. *Southern African Linguistics and Applied Language Studies*, *25*(1), 91–105. https://doi.org/10.2989/16073610709486448

Maddieson, I., & Sands, B. (2019). The sounds of the Bantu languages. In M. Van de Velde, K. Bostoen, D. Nurse, & G. Philippson (Eds.), *The Bantu languages* (2nd ed., pp. 79–127). Routledge. https://doi.org/10.4324/9781315755946-3

Makalela, L. (2013). Black South African English on the radio. *World Englishes*, *32*(1), 93–107. https://doi.org/10.1111/weng.12007

Nelson, C. (2020). *The handbook of world Englishes* (2nd ed.). Wiley-Blackwell.

Norro, S. (2021). Namibian teachers' beliefs about medium of instruction and language education policy implementation. *Language Matters: Studies in the Languages of Southern Africa*, *52*(3), 45–71. https://doi.org/10.1080/10228195.2021.1951334

Norro, S. (2024). Namibian teachers' practices in a multilingual context. *International Journal of Multilingualism, 21*(1), 360–378. https://doi.org/10.1080/14790718.2022.2065280

Norro, S. (2022). Factors affecting language policy choices in the multilingual context of Namibia: English as the official language and medium of instruction. *Apples – Journal of Applied Language Studies, 16*(1), Article 1. https://doi.org/10.47862/apples.107212

Odden, D. (2015). Bantu phonology. In Oxford Handbooks Editorial Board (Ed.), *Oxford handbook topics in linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199935345.013.59

Peltola, M. S., Lintunen, P., & Tamminen, H. (2014). Advanced English learners benefit from explicit pronunciation teaching: An experiment with vowel duration and quality. *AFinLA-e: Soveltavan Kielitieteen Tutkimuksia*, 0(6), 86–98.

Roach, P. (2004). British English: Received pronunciation. *Journal of the International Phonetic Association*, *34*(02), 239–245. https://doi.org/10.1017/s0025100304001768

Schröder, A., Zähres, F., & Kautzsch, A. (2021). The phonetics of Namibian English. In A. Schröder (Ed.), *The dynamics of English in Namibia: Perspectives on an emerging variety* (pp. 111–133). John Benjamins.

Stell, G. (2022). 'De-ethnicization' in New Englishes: Perception and recognition of ethnicity in Namibian English. *Lingua*, *274*, 103355. https://doi.org/10.1016/j.lingua.2022.103355
