**Investigating Variation in the Pronunciation of Spoken and Sung Text:**

**A Kimbra and Lorde Case Study**

BA Thesis English Language and Culture, Utrecht University

D.R. Jordens

3839249

Supervised by: Dr. K. Sebregts

Second reader: Dr. N. de Haas

August 2020

Abstract

This study aimed to answer the question of whether there was variation between the pronunciation of speech and singing of New Zealand artists Kimbra and Lorde. By means of phonetic analysis of interviews and the artists’ music, it was found that when comparing singing and speech, there was a shift features from standard New Zealand English towards standard American English and Southern British English realizations. The cause for these differences is not immediately apparent. Instead, it is argued, it is a multi-faceted system of sociolinguistic motivations which may all be odds with each other. These include motivations for language variation such as genre appropriateness, acts of identity, audience and referee design, and discourse-based motivations based on topic and mode.

**Table of contents**

1. Introduction ………...………………………………………………………………... 1
	1. New Zealand English ……………………………………………………….. 7
		1. Consonants ……………………………………………………..... 8
		2. Monophthongs …………………………………………………… 9
		3. Diphthongs ……………………………………………………….. 11
		4. Auckland English ………………………………………………… 12
		5. The present study ………………………………………………… 12
2. Methodology ...……………………………………………………………...……….. 12
	1. Subjects ……………………………………………………........................... 13
		1. Kimbra ........................................................................................... 13
		2. Lorde .............................................................................................. 13
	2. Material ........................................................................................................... 14
	3. Analysis ........................................................................................................... 14
3. Results .......................................................................................................................... 16
	1. Kimbra ............................................................................................................. 16
	2. Lorde ............................................................................................................... 21
	3. Monophthongization and rhoticity .................................................................. 24
4. Discussion .................................................................................................................... 25
5. Conclusion .................................................................................................................... 30

References .......................................................................................................................... 33

**Investigating Variation in the Pronunciation of Spoken and Sung Text:**

**A Kimbra and Lorde Case Study**

1. **Introduction**

The sociolinguistics of singing has gained an increasing amount of academic interest since Trudgill’s (1983) landmark study. Many different genres and languages have been investigated, but the main focus has been on English pop and rock vocals (e.g. O’Hanlon, 2006; Simpson, 1999). While there is variation in every (socio)linguistic aspect, it is most apparent in phonetics and phonology. Intraspeaker shifts in pronunciation between speaking and singing have been recorded as early as the 1920s and were prevalent in genres such as jazz and crooning (Trudgill, 1983). With American rock-and-roll music becoming widespread in the 1950s, British pop singers’ pronunciation showed American English (AmE) influences. This can be observed in, for example, the oeuvres of bands in the 1960s. In his paper, Trudgill (1983) analyzed 1960s albums of the Beatles, who hail from Liverpool and speak a non-rhotic variety, and found that in their first album, *Please Please Me,* non-prevocalic /r/ was as pronounced almost half the time. Throughout their albums, rhoticity saw a downward trend, and on *Abbey Road* it had decreased to approximately 5% of utterances. He found these trends in different variables, such as the BATH vowel which was more often realized as [æ], but later as [a:], and in different bands such as the Rolling Stones. In the 1970s, British punk bands more obviously returned to a British variety in their singing. This trend remains a topic of interest, even in the next century (e.g. Beal, 2009; Coupland, 2009).

At the start of his study, Trudgill (1983) states; “There can be no doubt that singers are modifying their linguistic behavior for the purposes of singing.” He found that there are six features that those singers generally display, as an attempt to mimic AmE varieties: Intervocalic /t/ is realized as an alveolar tap and the BATH vowel (transcribed as [a:] in Trudgill (1983), but generally [ɑ:] in Southern British English) is realized as [æ] in Southern British English. It is important to note here that this BATH-TRAP merger already was already a feature of Northern English, including Liverpool where the Beatles are from. Furthermore, there is a high degree of rhoticity, the PRICE vowel is monophthongized to [aˑ], words such as *love* and *done* are realized closer to a [ɘ] than [ʌ], and the LOT vowel is generally unrounded ([ɑ:]). However, besides the question of how singers change their linguistic behavior, Trudgill confronts the question “why do singers modify their pronunciation in this way?” (1983, p. 143). He provides a main reason as well as supplementary motivations that may conflict with one another.

As the field was dominated by Americans at that time, this was the demographic British pop artists wished to identify with musically, which lead them to modify their pronunciation in their singing to resemble the model group (Trudgill, 1983). This is a basic interpretation of what Tabouret-Keller and Le Page (1985) refer to as acts of identity. This does not mean that these modifications are successful by default; members of the model group do not necessarily recognize the linguistic behavior as accurate emulation of their features. According to Le Page, the emulation of the model group is constrained by to what extent the individual(s) can identify the model group, the level of access to the model group to work out the model group’s linguistic behavior, and the ability of the individual(s) to modify their behavior. This is further inhibited by the strength of motivations for retaining another’s model versus retaining a sense of unique identity, and these motivations may be conflicting (Le Page as cited in Trudgill, 1983).

According to Trudgill (1983), the varieties he studied that were being (successfully or unsuccessfully) mimicked were mostly Southern US and African-American English, because most 20th century genres have Southern American and/or African-American roots. These genres have linguistic style and register associated with them and linguistic usage within these genres is inspired by what is deemed appropriate. This notion of genre appropriateness may be a factor in Trudgill’s (1983) results on the pronunciation of the Beatles’ over time; a change in pronunciation reflects a change in genre. In her sociolinguistic appraisal of Australian hip-hop music, O’Hanlon (2006) found that, similarly to Trudgill, genre appropriateness is an obviously relevant factor, but not a sufficient explanation on its own for the variation in singing pronunciation. It can create a conflicting motivation, however, which can lead to retention of features of the genre when there is a strong desire for a unique identity, or even a different model group. O’Hanlon (2006) found that while Australian pop and rock artists still tend to model their pronunciation on American features, Australian hip-hop artists for the large part reject this tendency. They not only used more Australian English phonological features, but also more Australian English lexical and grammatical items. However, there was still a set of culturally specific American hip-hop jargon present.

Another motivation that can cause conflict, or a change of model group, is the notion of prestige. According to Trudgill (1983), the (covert or overt) use of features that are valued positively by speakers can aid identification with a group those features are associated with. He explains this with the punk movement in the 1970s, which showed a further decrease in the use of American pronunciation features. Punk-rockers used non-standard, low-prestige forms in their singing that were associated with British working-class youth. The employment of these in their singing while they did not use these in their speech is, Trudgill argues, proof of covert prestige which served to aid identification with the British working class. However, while the identity of the working-class British identity was emphasized, the retention of American features indicates that there are conflicting pronunciation models. Nevertheless, this focus on British youth identity has persisted into later popular genres such as indie rock, with the early work of bands such as Arctic Monkeys (Beal, 2009).

Simpson’s (1999) study presents a reevaluation of the American pronunciation model first proposed by Trudgill (1983). In what he dubs the USA-5 model, Simpson deviates from Trudgill’s model by disregarding the vowel modification in words such as *love.* Furthermore, he claims that, in singing, “[t]he Southern British English long open vowel [ɑ:] in words like *dance*, *last* and *ask* has a shorter, more advanced realization close to [a]. In those

lexical environments where Northern British English accents also contain an

/ɑ:/, as in 'half, a similar change occurs” (p. 345). Additionally, by touching upon the highly constructed nature of popular music genres Simpson presents two more motivations for pronunciation variation. More so than Trudgill, he attends to the discourse mode as an additional motivator for language variation and suggests that the mode of singing announces a departure of an individual’s speaking variety. Furthermore, the field of discourse, i.e. topic and pragmatics of language events can drive language variation, and Simpson argues that singers develop a linguistic persona that is divergent from their normal speech patterns.

In the 1990s, the popular music production landscape became largely United States-based as it was dominated by a handful of multinational US-based record companies (Simpson, 1999), and even though the American pronunciation model presented by Trudgill and Simpson no longer has the influence it had in the 1950s, the pronunciation features were still widespread in singing (Simpson, 1999, p. 363). Over the next decades, however, the pop music industry saw significant changes. The industry saw increasing globalization, with an influx of K-pop and J-pop artists. Furthermore, digital streaming became normalized for music consumption, with the Swedish company Spotify becoming a music industry lynchpin. Consequently, it became increasingly possible for independent record labels to compete. Nevertheless, after record label EMI was absorbed by Universal Music Group and Sony BMG in 2011, those two and Warner Music Group became the majority shareholders in the industry (Nielsen SoundScan & Nielsen BDS, 2013). As of 2018, the combined share these “big three” American companies have in the global music industry is as high as 66.2%. (Watson, 2020).

These American companies, however, started to decentralize in the late 1980s and 1990s, and the advent of music-focused television networks emphasized the importance of music videos. Simpson (1999) and Frith (1988) state that local divisions of these companies have resulted in fragmentation of rock and pop music genres. This fragmentation is exacerbated by, they argue, homogenization of the music as a result of video clips “exhibiting ‘an extraordinary sameness’ of genre and performance” (Simpson, 1999, p. 362). This has resulted in lexicalization of a large number of subgenres in rock and pop music (Simpson, 1999). These sub-genres all have their own conventions according to which rock and pop artists act (Frith, 1988, as cited in Simpson, 1999). The impact of these sub-genres and their conventions, Simpson (1999) claims, is unclear. He offers Liam Gallagher, lead singer of the band Oasis, as an example for the sub-genre britpop. At the time, Gallagher spoke and sung in a working-class Mancunian English variety, and held a very positive attitude of his Manchester roots, but also exhibited AmE and Liverpool English influences in his singing of *Wonderwall*. According to Simpson (1999) and Durant (1984), a possible reason for this lies in the musical influence from previous generations of artists in the same genre, and inheritance of this genre by subsequent generations in different social situations (Durant, 1984, p. 120-121, as cited Simpson, 1999, p.364). Americanisms, Simpson argues, could then be “epiphenomena rather than true indices of sociolinguistic motivation”, and a remnant of a generation in which the influence of the American pronunciation model was more pronounced (Simpson, 1999, p. 363-364).

Consequently, there are many different entities that pop artists speaking any vernacular may, consciously or subconsciously, attend to. This attention to others can influence a speaker to modify their linguistic behavior. This is at the core of Bell’s theory of audience and referee design (1984; 1990), which presents an alternative motivation for language variation to linguistic acts of identity. Bell (1984) analyzed the speech of New Zealand radio news broadcasters who read the news on two or more different radio stations, and found that readings on radio stations with a higher status audience displayed more standard (in this case the standard was Received Pronunciation) forms than on stations with lower status listeners. His theory entails that an audience or referee exercises its influence on a speaker to the extent that the speaker converges to or diverges from what they perceive to be the prestige variety. The referee, or reference group, is a third party that is not physically present which influences the linguistic choices of the speaker as if the speaker is addressing the referee, as “[t]hey hold an umpiring role in the speaker’s consciousness” (Bell, 1984, p. 161). For a recording musician, this could take the shape of a particular group of (imagined) listeners, for example, such as the punk-rocker who imagines they will be heard by British working class youth but does not necessarily exhibit the prestige variety in his own speech, and modifies their pronunciation accordingly (Trudgill, 1983; Bell, 1984). Consequently, the purpose of referee design is “for your speech to put the audience in mind of a particular reference group” (Bell, 1984, p. 190). While this is reminiscent of acts of identity, as both theories describe language variation in different social situations, the motivation for this is different. In audience and referee design, the speaker’s modification of their linguistic behavior is based on people: who they are addressing or who can overhear them. In acts of identity, however, the speaker has agency, and language is used for the construction of their identity; the speaker chooses to modify their speech to resemble that of the social groups they wish to identify with.

In summary, language variation in English-sung popular music has historically shown a tendency towards Southern US and African-American varieties. This led to American pronunciations being dominant in pop music (Trudgill, 1983; Simpson, 1999), the influence of which has diminished over time but remains present and widespread. This variation can be motivated by the speaker’s wish to identify with a certain group or individual and what is perceived as appropriate language for a certain genre, but also by possibly conflicting motivations such as covert prestige, genre appropriateness, subject matter and mode of discourse. Furthermore, external social situations and referees may cause the speaker to converge to their idea of an invisible referent’s variety.

However, the academic literature in this field has predominantly featured male-fronted British bands and male artists. While Bell (1984) presented New Zealand as a particularly interesting area for studies of sociolinguistics of media, the attention New Zealand English (NZE) has received in this field is limited to different modes of discourse (e.g. Bell, 1984), and, in the case of popular music, to locally known male artists (Gibson, 2005; 2010). In the latter, the male artists were shown to converge to American norms, even when they held positive views about NZE singing. Even though these artists rejected features that were considered stereotypically American, the American model still resonates in New Zealand pop music. However, “if a range of high-profile artists started using NZE in their singing, this situation could change quickly” (Gibson, 2010, p. 166).

The years that followed Gibson’s (2010) study saw the rise of high profile, female New Zealand singers with international success such as Lorde and Kimbra, who come from Auckland and Hamilton respectively. A more detailed profile of these singers is presented in the following chapter. By analyzing the phonetics and phonology of their singing, the aim of the present study is to provide an insight into the progression of the sociolinguistic situation of NZE pop music since Gibson’s (2010) study. To be able to compare the artists’ speech with their singing pronunciation, the remainder of this introductory chapter will discuss NZE phonology.

* 1. *New Zealand English*

This study is an investigation of New Zealand English (NZE) spoken and sung by New Zealanders of European descent (Pākehā) (Bauer et al., 2007). An overview of the phonemic inventory of NZE spoken by educated, middle class women is adapted from Bauer and Warren (2004) and Bauer et al. (2007). It is important to note that this inventory presents an extensive but generalized view of phonemes and does not necessarily account for idiolects or regional variation. Bauer et al. (2007) claim that "[m]uch of this variation (...) is lost in the formal environment of reading a short passage for a microphone" (2007, p. 97), which is the type of data Bauer et al. (2007) used.

* + 1. *Consonants*

To discuss differences, Bauer et al. (2007) refer to RP. As, for example, the DRESS vowel is raised in NZE but also in traditional RP, this paper will refer to Standard Southern British English (SSBE) instead to disambiguate. The phonemic inventory for NZE consonants can be found in Table 1 below.

*Table 1.* *Consonants of New Zealand English*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Bilabial | Labio-dental | Dental | Alveolar | Post-alveolar | Palatal | Velar | Glottal |
| Plosive | p b |  |  | t d |  |  | k ɡ |  |
| Affricate |  |  |  |  | ʧ ʤ |  |  |  |
| Fricative |  | f v | θ ð |  | ʃ ʒ |  |  | h |
| Nasal |  m |  |  | n |  |  | ŋ |  |
| Approximant |  w |  |  | ɹ |  | j | w |  |
| Lateral approximant |  |  |  | l |  |  |  |  |

NZE consonants are similar to those in other English varieties. For the purpose of this study, it is important that intervocalic /t/ is generally realized as a voiced tap in NZE (Bauer et al, 2007). As this is also a feature of the American pronunciation model (Trudgill, 1983; Simpson, 1999), this variable will be disregarded in the analysis of the present study. In contrast, NZE is a non-rhotic variety. It has some exceptions, however. Bauer & Warren (2004) found that the name of the consonant *r* and *Ireland* are always uttered with a non-prevocalic [ɹ], and that there is a degree of variable rhoticity particularly following the NURSE vowel (also found in Gibson, 2005). Additionally, there is evidence of lexically-driven, pseudo-American realizations of /r/ in expressions borrowed from American media, such as *whatever*.

* + 1. *Monophthongs*

*Table 2. Lexical sets of monophthongs and their realizations in NZE.*

|  |  |
| --- | --- |
| **Stressed** |  |
| FLEECE | iː, (ɪɨ) |
| BATH, START, PALM | ɐː |
| NURSE | ɵː, (**œː** øː) |
| THOUGHT, NORTH, FORCE | oː, (oə, o.ɐ) |
| GOOSE | ʉː, (ʏː, ɪʉ, ɐʉ) |
| KIT | ɘ, (ə, ə̝, ɪ) |
| DRESS | e, (e̝, e̝̠) |
| TRAP | ɛ |
| STRUT | ɐ, (ɐ̟) |
| LOT | ɒ, (ɞ̠) |
| FOOT | ʊ, (ɨ̞) |
| FACE | æe, (ɐe, ɐi) |
| PRICE | ɑe, (ɒe, ɑi) |
| CHOICE | oe, (oi) |
| GOAT | ɐʉ, (ɐɨ) |
| MOUTH | æo, (æʉ, ɛʉ) |
| NEAR | iə, (iː.ɐ, e.ɐ, eə) |
| SQUARE | eə, (iː.ɐ, e.ɐ, iə) |
| CURE | ʉə, (ʉ.ɐ) |
| **Unstressed** |  |
| commA, horsES | ɘ, (ə, ɜ, ɐ) |
| happy | i, (ə̯i, iˑ) |
| treacLE | ɯ, (ɔ̝, o, ʊ, ɤ) |

Table 2 above shows the vowel inventory for NZE, modelled after Bauer & Warren (2004) and Bauer et al. (2007). These vowels are presented as realizations of the vowels in the lexical sets of Wells (1982). It is important to note that the sounds which precede the bracketed sounds are, in the terminology of Bauer et al. (2007), generalized, and are used for (broad) phonemic transcription. The bracketed sounds have been found to be alternative realizations for these lexical sets. Some of the variation is allophonic and lexically driven, as the first vowel in *proven*, for example, “is often pronounced with GOAT in the stressed syllable as an alternative to GOOSE” (Bauer & Warren, 2004, p. 600). A lot of the variation, however, is regional or sociolinguistic.

Most notably, the KIT vowel is centralized in NZE, so much so that there is nearly no distinction between the KIT and commA vowels. However, it is generally perceived as closer to an [i] before [ŋ] and, to a lesser extent, [ɡ], such as in *sing* or *big.* The DRESS vowel is very close in NZE, and resembles the SSBE KIT vowel, while the NZE TRAP vowel is more akin to the SSBE DRESS vowel. Furthermore, Bauer & Warren (2004) and Warren (2018) discuss that lip-rounding is always weak in NZE. The fact that the NURSE this vowel is generally rounded ([ɵː]) in NZE, seemingly regardless of whether the lip is rounded or unrounded, then, indicates that there is an as of yet unknown underlying articulatory mechanism that gives NZE vowels a rounded quality without use of the articulatory gesture (Bauer & Warren, 2004; Warren, 2018). Furthermore, NZE contains considerable neutralization before /l/ and /r/. Although there is variation in neutralization in these environments, before /r/ FLEECE and NEAR generally neutralize to [iː], DRESS and SQUARE to [e], and GOOSE and CURE to [ʉː]. It is important to note that this is only the case in disyllabic contexts, as NZE is a non-rhotic variety. Sounds preceding /l/ show more variation. KIT, FOOT, GOOSE and STRUT vowels are often neutralized, and as a result *pill, pull* and *pool,* as well as *cult* and *kilt* may be virtually indistinguishable. Moreover, neutralized LOT, GOAT, THOUGHT and FOOT vowels can lead to homophony between *poll, pole, pull* and *Paul* (Bauer et al., 2007).

* + 1. *Diphthongs*

The FACE vowel is generally more open, while the PRICE vowel is further back than their RP counterparts. Consequently, the latter may be perceived as a SSBE CHOICE diphthong. Furthermore, the starting point of the NZE CHOICE diphthong is more akin to the NZE THOUGHT vowel. MOUTH diphthongs have a close starting point, and are seemingly homophonic to an SSBE TRAP vowel plus /l/ such as in *help.* Bauer & Warren (2004) found that NEAR and SQUARE were being increasingly merged, especially by younger speakers. Finally, the CURE vowel only occurs after /j/, and words like *poor* and *tour* contain FORCE vowels. “The overall result is that the CURE vowel has very little functional load in New Zealand English” (Bauer et al., 2004)

* + 1. *Auckland English*

One of the artists discussed in this study, Lorde, is a young speaker from Auckland, New Zealand. In her thesis, Ross (2018) presents an analysis of Auckland (New Zealand) English vowels and found significant age effects for the realizational variation of certain vowels. For example, where the FLEECE usually had a vowel onglide to resemble a diphthong, for young Auckland speakers this onglide decreased to the point that it was generally realized as a monophthong. Furthermore, the DRESS and TRAP vowels that are generally raised in NZE, are in the process of becoming lowered for young Auckland speakers. For the same group, the KIT vowel is fronted. In terms of diphthongs, FACE is markedly raised and the span between start and finish point is smaller. Additionally, PRICE was found to be fronted, and the starting point of GOAT was raised (Ross, 2018). These findings are particularly interesting as they indicate departures from salient NZE towards standard realizations in Standard Southern British English (SSBE) and AmE.

* + 1. *The present study*

As mentioned before, the purpose of the present study is to provide an analysis of more recent language use in New Zealand pop music. It aims to provide insight into the development of the American pronunciation model in the years after Gibson’s (2010) study, which saw the rise of high-profile New Zealand pop music artists Kimbra and Lorde. As previous research in this field has generally focused on male singers of different English varieties and age groups, and on different modes of discourse, it may be important to investigate the language use of young Pākehā women to complement the existing data. The research question for this study is, therefore: Do Kimbra and Lorde modify their respective English varieties when singing compared to speaking? In Trudgill (1983) and subsequent research in the same field on both British English (Simpson, 1999, Beal, 2009), Australian English (O’Hanlon, 2006), and New Zealand English (Gibson, 2010), the American pronunciation model is shown to be present, especially in pop genres. This was the case even when artists overtly rejected the American model in favor of another model, as exemplified by Punk-rock music. Because the American model has persisted through overt rejection, music industry globalization and de-centralization, the expectation is that both Lorde and Kimbra would exhibit at least some degree of style-shifting towards American varieties of English. Any findings in this study will be discussed in the light of the theories discussed in the first chapter.

1. **Methodology**

The methodology for this study was borrowed from Trudgill (1983) and Gibson (2010), and was inspired by Bell’s (1984) work on media analysis of New Zealand English. The main differences between the American pronunciation model used in Trudgill (1983) and Simpson (1999) and NZE are the realizations of certain vowels and rhoticity. These are therefore the variables investigated in this study. Trudgill (1983) presented a quantitative analysis of the phonetics and phonology of the singing of popular bands in the 1950s and 1960s, and Gibson (2010) explored the works by New Zealand artists in a similar way. Their analysis was enriched with analysis of linguistic features of grammar, semantics and pragmatics in the lyrics. Gibson presented his results in the form of formant tables and vowel spaces. Consequently, the present study features a similar quantitative analysis of materials provided by Lorde and Kimbra, and uses the same medium of presentation for results.

* 1. *Subjects*
		1. *Kimbra*

Kimbra Johnson is a singer from Hamilton whose music is a blend of (art) pop and various other genres, including jazz and rock, with influences from artists such as Prince, Minnie Riperton, Björk and Jeff Buckley. Before releasing her debut album *Vows* (Johnson, 2011) Kimbra had moved to Melbourne. She recorded the album between 2008 and 2011 here with Warner Bros. Records, and the album went platinum in New Zealand and Australia. It featured singles such as “Settle Down” and “Cameo Lover”. In 2012, she gained popularity as a result of her collaboration with Australian singer Gotye in “Somebody that I Used to Know”.This song won her two Grammy awards in 2013, for“Record of the Year”and “Best Pop Duo/Group Performance”*.* Afterwards, she moved to Los Angeles to work on her second album, *The Golden Echo*, and released her latest album, *Primal Heart*,in 2018.

* + 1. *Lorde*

Ella Yelich-O’Connor, also known as Lorde, is a Croatian-New Zealand singer-songwriter from Auckland. Her music blends various genres of pop music with electronic elements, and while her early music drew from hip-hop and minimalism, her later work experimented with heavier instrumentation and maximalism. She signed with Universal Music Group in 2009 as a developing artist and teamed up with writer and producer Joel Little, known for working with Taylor Swift, Sam Smith, and Imagine Dragons. Lorde self-released her debut “The Love Club EP” in 2012 and its success caused UMG to commercially release it in 2013. The EP contained the single “Royals”, which was a commercial success internationally. Her certified multi-platinum album *Pure Heroine* was released in 2013.

* 1. *Material*

Data of Lorde’s speech was drawn from multiple interviews (Lorde Updates, 2013; Newshub, 2015) conducted by New Zealand media sources to get the most accurate representation of her spoken idiolect and eliminate potential confounds with regards to speech accommodation towards non-NZE speakers (Giles & Powesland, 1997). It is important to note that Lorde expresses a positive attitude towards New Zealand, as she calls it her home and describes how she looks forward to returning here. Her album *Pure Heroine* (Lorde, 2013) was the subject of analysis regarding her sung pronunciation.

For Kimbra, spoken data was collected from three interviews (Kimbramusic, 2011; Timt4stic, 2011; Williams, 2011), which were chosen because they occurred in New Zealand, before she had moved to the US. There were no interviews available, however, that were conducted before she moved to Australia. This could result in Australian English features in the data. She does, however, explicitly express a positive attitude towards her home country in the interviews, even comparing it to Australia. Songs from her album *Vows* (Johnson, 2011) were used for analysis of her sung pronunciation.

* 1. *Analysis*

Phonetic analysis for this study was performed using the computer program Praat (Boersma & Weenink, 2019). The Audio files that were analyzed were extracted from YouTube and converted to mp3, then loaded in Praat, and the audio files of interviews were then manually edited to contain only the parts spoken by the artists. The audio files were manually edited. For spoken and sung vowels the first three formants were recorded in a separate Microsoft Excel file, and rhotics were counted as a Labovian discrete variable, as done in Trudgill (1983). Monophthongs were measured at the midpoint where formants were generally most stable and diphthongs were measured closest to both its starting point and ending point, while making sure there was no coarticulation. The discrete rhoticity variable was scored as a percentage of the total occurrences of non-prevocalic /r/ realized, in a different sheet in the same file. Data for vowels in that Excel file was phonologically arranged according to Wells’ (1982) lexical sets.

During the analysis of the music, there were multiple occurrences of where Praat could not calculate the formants of the lead vocals due to musical backing or the presence of prominent backing vocals. To make analysis for these segments possible, the mp3 file was loaded into sound recording and editing program Audacity (Audacity Team, 2020) and the vocal reduction and isolation effect was used to isolate the lead vocals. While the solo vocal track this generated was diminished in quality, most of the time clear formants could be recorded. The presence of music and backing vocals in these instances, as well as the diminished quality in the files with isolated vocal could have, to some extent, made the measurements less accurate and reliable. However, there was no reason to assume the data was unusable, and in the cases this method did yield unusable results, the relevant segments were not recorded.

IBM SPSS Statistics 25 was used for testing the data for significant differences in formant frequencies between conditions. In his paper, Gibson (2010) used the Wilcoxon signed-rank test to indicate significant differences between sung and spoken formant frequencies based on the median of the observations, as he found that there were discrepancies between the first formants of sung versus spoken vowels. The present study also found these discrepancies in multiple vowels and employed the same test account for significant differences. This test was chosen over paired-sample t-tests due to smaller sample sizes for some vowels. The smaller sample sizes frequently prevented an accurate indication of normally distributed data and increased the risk of the mean being disproportionately affected by outliers. Furthermore, the data in this study was skewed more often than not, which favors the assumption that the data is generally not normally distributed and therefore the Wilcoxon signed-rank test is employed. This test assumes a null-hypothesis and is therefore generally regarded as two-tailed. However, Gibson (2010) found that there was a tendency for sung F1 to be higher. Therefore, a case could be made for paired-sample differences with p < 0.1 reaching significance in a one-tailed approach. These will be highlighted in the present study, but as caution is advised for the interpretation of these results this section will focus on the unambiguously significant differences.

1. **Results**
	1. *Kimbra*

The results for Kimbra are plotted in Table 3 below. For visual clarity, Figure 1 and Figure 2 feature vowel spaces based on mean F1 and F2 for monophthongs and diphthongs in both speaking and singing conditions. For reference on NZE, the discussion of results will generally reference the data and vowel spaces gathered from three female speakers in Bauer et al. (2007), two of which were young MA students in their twenties and the other was older, complemented with the data from female speakers in Bauer & Warren (2004). At a glance, Kimbra has a larger spread in her vowel space than Lorde (Figures 3 and 4) in terms of the open-closed dimension, especially in more central vowels. This dimension has an inverse

*Table 3. Mean (total observations) first and second formant frequencies in Hz per condition, the difference between conditions in Hz, and the result of the Wilcoxon test for Kimbra. For monophthongs, measurements were taken at the midpoint. Diphthongs were measured at starting and ending points while taking coarticulation into account.*

|  |  |  |  |
| --- | --- | --- | --- |
| Vowel | F1 |  | F2 |
|   | sing (n) | speak (n) | diff. | sig. |  | sing | speak | diff. | sig. |
| FLEECE | 465 (25) | 439 (44) | 26 | n.s. |  | 2378 | 2332 | 46 | n.s. |
| START | 974 (17) | 770 (8) | 205 | \*\* |  | 1470 | 1346 | 124 | \*\* |
| NURSE | 712 (11) | 560 (11) | 152 | \*\*\* |  | 1736 | 1411 | 325 | \*\*\* |
| FORCE | 723 (17) | 489 (16) | 235 | \*\*\*\* |  | 1277 | 1014 | 263 | \*\*\* |
| GOOSE | 445 (32) | 397 (21) | 48 | \*\*\*\* |  | 1529 | 1759 | -230 | \*\*\* |
| KIT | 589 (31) | 515 (77) | 76 | \*\*\*\* |  | 2070 | 1760 | 310 | \*\*\*\* |
| DRESS | 745 (24) | 524 (22) | 221 | \*\*\*\* |  | 2053 | 2007 | 46 | n.s. |
| TRAP | 761 (14) | 669 (22) | 92 | \*\* |  | 1974 | 1820 | 154 | n.s. |
| STRUT | 754 (21) | 635 (26) | 119 | \*\*\* |  | 1471 | 1375 | 96 | n.s. |
| LOT | 714 (29) | 604 (21) | 110 | n.s. |  | 1351 | 1213 | 138 | n.s. |
| FOOT | 577 (10) | 449 (5) | 128 | n.s. |  | 1439 | 1543 | -104 | n.s. |
| commA | 595 (36) | 488 (58) | 107 | \*\*\*\* |  | 1737 | 1594 | 143 | \*\*\* |
| happY | 450 (12) | 447 (5) | 3 | n.s. |  | 2320 | 2210 | 110 | n.s. |
|  |  |  |  |  |  |  |  |  |  |
| FACE (a) | 731 (21) | 581 (13) | 150 | \*\*\* |  | 2105 | 1914 | 191 | \*\*\* |
| FACE (b) | 484 | 512 | -28 | \*\*\* |  | 2317 | 2138 | 179 | n.s. |
| PRICE (a) | 872 (19) | 729 (23) | 143 | \*\* |  | 1649 | 1513 | 136 | n.s. |
| PRICE (b) | 654 | 542 | 112 | n.s. |  | 2044 | 1837 | 207 | \*\* |
| CHOICE (a) | 501 (1) | 528 (2) | -27 | N/A |  | 1494 | 1230 | 264 | N/A |
| CHOICE (b) | 570 | 485 | 85 | N/A |  | 2415 | 1883 | 532 | N/A |
| GOAT (a) | 742 (22) | 562 (18) | 180 | \*\*\* |  | 1321 | 1569 | -248 | \* |
| GOAT (b) | 491 | 456 | 35 | n.s. |  | 1241 | 1473 | -232 | \*\* |
| MOUTH (a) | 699 (18) | 774 (12) | -75 | \* |  | 1855 | 1591 | 264 | \*\*\* |
| MOUTH (b) | 510 | 613 | -103 | \*\* |  | 1298 | 1325 | -27 | n.s. |
| NEAR (a) | 516 (3) | 472 (6) | 44 | N/A |  | 2632 | 2237 | 395 | N/A |
| NEAR (b) | 604 | 523 | 81 | N/A |  | 1842 | 1813 | 29 | N/A |
| SQUARE (a) | 647 (5) | 525 (6) | 122 | N/A |  | 2191 | 2084 | 127 | N/A |
| SQUARE (b) | 691 | 516 | 175 | N/A |  | 1967 | 1840 | 0 | N/A |
|  | \* for p < 0.10, \*\* for p < 0.05, \*\*\* for p < 0.01, \*\*\*\* for p .001; N/A for too small sample size |

 *Figure 1. Vowel space for Kimbra’s monophthongs.*

 *Figure 2. Vowel space for Kimbra’s diphthongs*

correlation with F1 frequency; higher F1 generally indicates a lower, or more open, vowel. The front-back dimension in turn corresponds with F2 frequency. Kimbra’s FLEECE and happY vowels tend to have a higher F1 when compared to the female NZE speakers in Bauer et al. (2007) and Bauer & Warren (2004), which underpins her tendency to pronounce some sounds with a more open vocal tract setting. The BATH/START/PALM, and STRUT vowels, which, instead of being the open, centralized vowel they tend to be in NZE, gravitate towards the NZE LOT position in terms of formants. This is compatible with lower front vowels, as the vowel system is shifted down and back. The LOT vowel raises and backs as well and seems to resemble [ɔ], which is considered a standard realization for varieties of NZE. Furthermore, NURSE has a lower F2, and thus is further back than in standard NZE. Similarly, the commA vowel also finds itself somewhat further back and raised, but still falls in the NZE range according to Bauer et al. (2007).

Most of Kimbra’s spoken monophthongs can be considered standard NZE, however. While the FLEECE and happY vowels have a slightly raised F1, they still have a close and front position. Her FORCE vowel can be found at the close-mid (almost near-close) position, while FOOT is firmly in the NZE-standard close-mid, central-back position. The salient KIT, DRESS and TRAP vowels are centralized, raised and fronted, and raised, respectively, in Kimbra’s speech as well. Furthermore, the GOOSE vowel is fronted and resembles an [ʉ].

Kimbra’s diphthongs are generally very similar to the NZE diphthongs found by Bauer et al. (2007). It is important to note that there were only a small number of tokens for the CHOICE, NEAR, and SQUARE diphthongs, and discussion of these vowels is meant as a point of interest rather than an assertion. Compared to the data from Bauer et al. (2007) the trajectory of all Kimbra’s diphthongs is largely the same, but for some the span is smaller. This can be seen in the PRICE and CHOICE versus NEAR and SQUARE diphthongs in particular. In this study’s data, the finish targets for these diphthongs cluster together around (a slightly raised) [ə]. This is normal for NEAR and SQUARE. In contrast, the PRICE and CHOICE diphthongs tend to finish around a close-mid fronted position, and thus their finish displays a slightly lowered mean F2 than is expected for NZE (Bauer et al., 2007). Furthermore, instead of the conventional [æ] starting point, the starting point of the FACE vowel is raised to where it resembles an [eɪ] glide and the starting point of MOUTH is backed towards the [ɐ] territory. The highest degree of deviation from Bauer’s diphthongs is found in the GOAT vowel, however, as the starting point is markedly raised and rather than ending at [ʉ], the diphthong approximates [əʊ].

 For the comparison of speech and singing, Table 3 and Figures 1 and 2 offer valuable insight into the differences between Kimbra’s pronunciation of spoken and sung text. Most notably, while singing, the lowered DRESS, TRAP, and fronted KIT vowels indicate a departure from their typical NZE realizations and move towards standard AmE or SSBE ones. These differences are not only visual, but supported by the significance of the Wilcoxon tests. The latter is not the case for the LOT vowel, however, though it looks like it moves away from NZE similarly to the other vowels. The large number of positive F1 differences for the singing condition could indicate that this mode employs a more open upper vocal tract setting, or it could be an automatic effect of raised F0 in the singing mode, both of which are in line with Gibson’s (2010) findings. This also goes for the start position of the FACE. PRICE and GOAT diphthongs. The raised F1, when paired with F2 differences, cause some lexical sets to fall into a different allophone space. FORCE/NORTH/THOUGH, LOT, and STRUT cluster around [ɒ]. NURSE when sung is open-mid and centralized to approximate [ɜː]. In contrast, the START vowel is closer to Bauer et al.’s (2007) when sung.

The differences between diphthong glides are increased in the singing condition. When looking at the significant differences between conditions the closing diphthongs with the exception of MOUTH have a lower starting point, the start and finish of which are both slightly raised. The higher F2 also causes Kimbra’s sung FACE, PRICE and MOUTH diphthongs to become more like the vowels found by Bauer et al. (2007). It is the GOAT vowel, however, that deviates further from NZE by virtue of decreased mean F2. There were not enough data points to determine whether there were significant differences for the other vowels, though the same tendency of raised F1 and F2 can also be observed here. Additional data for these sets should point out whether this is truly the case.

* 1. *Lorde*

The results for Lorde’s phonetic analysis are plotted in Table 4, and vowel spaces can be found in Figures 3 and 4 below. Lorde’s monophthongs largely display the same tendencies as those seen in Kimbra’s data. In her speech, the DRESS and TRAP vowels are raised, while the KIT vowel and GOOSE vowels are centralized. Lorde’s LOT vowel is even more raised than Kimbra’s and is close to merging with the FORCE vowel when spoken. Similar to Kimbra, the STRUT and NURSE vowels are markedly raised and backed, respectively, when compared to NZE. In contrast, the apparent cluster of the KIT, commA and FOOT vowels is indicative of NZE, which displays a degree of overlap between these lexical sets. Lorde’s spoken diphthongs also closely resemble standard NZE diphthongs. Some notable differences can be found in the lower starting position for NEAR and, like Kimbra, the raised starting position for FACE and GOAT. The latter, also like Kimbra, ends further in the back than the conventional NZE GOAT diphthong. When singing, Lorde displays the same tendency as Kimbra. It is clear that, as FLEECE, happY and central NURSE do not move, there is a shift towards standard AmE or SSBE or AmE in the DRESS, KIT and TRAP vowels. This time, there is also a significant lowering in the LOT vowel as well, which could indicate that this might not be coincidental for Kimbra either. Additionally, GOOSE deviates from NZE as well the vowel is significantly backed towards a more open version of [u:], indicative of an intermediate form between NZE and AmE or SSBE. Lorde’s START vowel, however, is in

*Table 4. Mean (total observations) first and second formant frequencies in Hz per condition, the difference between conditions in Hz, and the result of the Wilcoxon test for Lorde. For monophthongs, measurements were taken at the midpoint. Diphthongs were measured at starting and ending points while taking coarticulation into account.*

|  |  |  |  |
| --- | --- | --- | --- |
| Vowel | F1 |   | F2 |
|   | sing (n) | speak (n) | diff. | sig. |   | sing (n) | speak (n) | diff. | sig. |
| FLEECE | 422 (39) | 382 (28) | 40 | \* |  | 2446 | 2273 | 173 | \*\* |
| START | 798 (9) | 890 (18) | -92 | n.s. |  | 1614 | 1514 | 100 | n.s. |
| NURSE | 605 (2) | 571 (18) | 34 | n.s. |  | 1612 | 1563 | 49 | n.s. |
| FORCE | 610 (11) | 474 (19) | 136 | \*\*\* |  | 1247 | 1046 | 201 | \*\*\* |
| GOOSE | 471 (19) | 378 (33) | 93 | \*\* |  | 1412 | 1749 | -337 | \* |
| KIT | 565 (71) | 450 (50) | 115 | \*\*\*\* |  | 2073 | 1764 | 309 | \*\*\*\* |
| DRESS | 648 (35) | 454 (28) | 196 | \*\*\*\* |  | 2061 | 1996 | 65 | n.s. |
| TRAP | 750 (27) | 562 (36) | 188 | \*\*\* |  | 1959 | 1871 | 88 | \*\* |
| STRUT | 784 (28) | 606 (24) | 178 | \*\*\* |  | 1573 | 1448 | 125 | \*\* |
| LOT | 751 (26) | 529 (38) | 222 | \*\*\*\* |  | 1345 | 1186 | 159 | \*\* |
| FOOT | 515 (3) | 455 (13) | 60 | n.s. |  | 1658 | 1547 | 111 | n.s. |
| commA | 556 (74) | 443 (49) | 113 | \*\*\*\* |  | 1792 | 1639 | 153 | \*\*\* |
| happY | 421 (11) | 355 (15) | 66 | n.s. |  | 2497 | 2235 | 262 | \*\*\* |
|  |  |  |  |  |  |  |  |  |  |
| FACE (a) | 594 (21) | 503 (20) | 91 | \*\*\* |  | 2092 | 1826 | 266 | \*\*\* |
| FACE (b) | 499 | 358 | 141 | \*\*\*\* |  | 2465 | 2140 | 325 | \*\*\* |
| PRICE (a) | 894 (33) | 595 (45) | 299 | \*\*\*\* |  | 1620 | 1544 | 76 | n.s. |
| PRICE (b) | 702 | 420 | 282 | \*\*\* |  | 1886 | 2035 | -149 | \*\* |
| CHOICE (a) | 723 (3) | 426 (9) | 297 | N/A |  | 1363 | 930 | 433 | N/A |
| CHOICE (b) | 623 | 401 | 222 | N/A |  | 2220 | 2028 | 192 | N/A |
| GOAT (a) | 637 (19) | 521 (32) | 116 | \*\*\* |  | 1412 | 1529 | -117 | \*\* |
| GOAT (b) | 439 | 387 | 52 | n.s. |  | 1246 | 1421 | -175 | \*\* |
| MOUTH (a) | 932 (17) | 687 (14) | 245 | \*\* |  | 1725 | 1701 | 24 | n.s. |
| MOUTH (b) | 615 | 508 | 107 | n.s. |  | 1502 | 1464 | 38 | n.s. |
| NEAR (a) | 760 (4) | 566 (5) | 194 | N/A |  | 2226 | 2127 | 99 | N/A |
| NEAR (b) | 654 | 628 | 26 | N/A |  | 1848 | 1891 | -43 | N/A |
| SQUARE (a) | 618 (6) | 441 (7) | 177 | N/A |  | 2316 | 1987 | 329 | N/A |
| SQUARE (b) | 647 | 511 | 136 | N/A |   | 1795 | 1782 | 13 | N/A |
| \* for p < 0.10, \*\* for p < 0.05, \*\*\* for p < 0.01, \*\*\*\* for p < 0.001 |  |  |  |  |  |

*Figure 3. Vowel space for Lorde’s monophthongs.*

 *Figure 4. Vowel space for Lorde’s diphthongs.*

the central open position in both conditions and the STRUT vowel assumes a more NZE-like position as well. Much like Kimbra’s data showed, FACE, PRICE, GOAT and MOUTH diphthongs significantly differ from the spoken condition as the higher F1 tendency is also observed here in every case. Where Kimbra’s MOUTH vowel was raised, Lorde’s significant F1 deviations were all positive, and thus lowered. In terms of F2 differences, Lorde’s GOAT vowel, too, was backed in the singing condition. Unlike Kimbra, however, she shows a significantly backed end point for PRICE.

* 1. *Monophthongization and rhoticity*

The diphthongs were tested for any signs of monophthongization. There were a few occurrences of, for example, *I* [aɪ] being uttered as [aː]. By comparing the F1 and F2 values respectively by means of Wilcoxon signed-rank tests, significant differences could be determined between the first and second sound in the diphthongs. The results are summarized in Table 5 below. The data presented no evidence for systematic monophthongization. The only

 *Table 5. Results of Wilcoxon tests for significant differences between F1 in start and ending position for diphthongs.*

|  |  |  |  |
| --- | --- | --- | --- |
|   | Kimbra |   | Lorde |
|  | Singing | Speaking |  | Singing | Speaking |
|   | F1 | F2 | F1 | F2 |   | F1 | F2 | F1 | F2 |
| FACE | \*\*\* | \*\* | \*\* | \*\* |  | \*\* | \*\* | \*\*\* | \*\*\* |
| PRICE | \*\*\* | \*\* | \*\*\* | \*\*\* |  | \*\*\* | \*\*\* | \*\*\* | \*\*\* |
| CHOICE | N/A | N/A | N/A | N/A |  | N/A | N/A | \*\* | \*\* |
| GOAT | \*\*\* | \* | \*\* | n.s. |  | \*\*\* | \* | \*\*\* | \* |
| MOUTH | \*\*\* | \*\* | \* | \*\* |  | \*\* | \*\* | \*\* | \*\* |
| NEAR | N/A | N/A | N/A | N/A |  | N/A | N/A | N/A | N/A |
| SQUARE | n.s. | \* | n.s. | \* |   | n.s. | \* | \* | \* |
| \* for p < .05, \*\* for p < .01, \*\*\* for p < .001, N/A for small sample size |  |  |  |

places that did not significantly shift were the GOAT ending point in Kimbra’s speech, and almost all starting points for the SQUARE diphthong in every condition.

*Table 6. Percentages of non-prevocalic /r/ realized per condition.*

|  |  |  |
| --- | --- | --- |
|   | Spoken | Sung |
| Kimbra | 15 | 27 |
| Lorde | 11 | 34 |

Finally, the results for the measurements of rhoticity are plotted in Table 6 above. As discussed in the method section, these were measured as the share of total non-prevocalic /r/ realized. There were a total of 199 spoken non-prevocalic /r/ occurrences (n = 124 for Kimbra, n = 75 for Lorde), and a total of 145 in the singing condition (n = 84 for Kimbra, n = 61 for Lorde) The results are quite similar; both artists’ variety is not complete non-rhotic when spoken. As stated in section 1.1.1., Bauer & Warren (2004) argue that there is some degree of variable rhoticity, particularly after a NURSE vowel, in NZE. This could be an explanation for the rhoticity found in the speech condition. However, the use of non-prevocalic [ɹ] increases while singing. Lorde’s “Royals” in particular displayed a relatively high degree of rhoticity, with 50% of all non-prevocalic /r/ being realized in that song. The extent to which non-prevocalic /r/ is realized could be evidence of an AmE pronunciation model present in singing pronunciation.

1. **Discussion**

The purpose of this study was to investigate whether Kimbra and Lorde modify their pronunciation when singing compared to speaking. Based on the data from this study, it can be inferred that this is indeed the case.

First, the artists’ speech was compared to that of other NZE speakers. It is important to reiterate that in the interviews investigated in this study revealed that both speakers were explicit in their positive attitudes towards New Zealand. For Kimbra, the data found is generally in line with that of Bauer et al. (2007) in terms of speech. Some vowels deviate somewhat from their data, but can be accounted for with the NZE data from Bauer & Warren (2004). Furthermore, Kimbra’s diphthongs also generally corroborated the data presented in Bauer et al. (2007) with some notable exceptions. The starting points for the FACE and GOAT vowel, for example, are raised and resemble typical SSBE (or AmE in the case of FACE) realizations.

A possible explanation is that the raising of the starting point of GOAT vowels that is found in young speakers of Auckland English (Ross, 2018) is also present in the variety spoken in Hamilton, a city close to Auckland where Kimbra is from. Another explanation is proposed by Gordon & Maclagan (2004), who discuss the concept of Cultivated, General, and Broad NZE. This division is based on social class, where members of a higher social class will speak more Cultivated NZE and will avoid using features of the General and Broad variety of lower social classes. Cultivated NZE is generally associated with educated women, and its features often closely resemble the features of Received Pronunciation, which, for example, also displayed the raised TRAP vowel. This argument already touches upon the sociolinguistic motivation of overt and covert prestige that can conflict with other motivations such as group identity. In this case, the overt prestige associated by a social group with some SSBE features may impinge on standard NZE, and two pronunciation models may be in conflict with each other as a result. However, the fact that only the GOAT and FACE diphthongs demonstrate a departure from General NZE and there is no single explanation for this phenomenon could warrant a closer look into the status of these diphthongs in young female speakers of NZE.

 To some extent, the same explanation could be relevant for Lorde’s speech specifically. Lorde’s STRUT and LOT vowel could be evidence of this as they are realized closer to [ʌ] and [ɔ], similar to SSBE, respectively. Furthermore, the starting points for the GOAT and FACE vowels are raised, but like for Kimbra, the explanation of the influence of Cultivated NZE or the findings by Ross (2018) remains unconvincing. For example, Lorde’s DRESS, TRAP and KIT vowels do not follow the process of changing found by Ross (2018) and instead are in the expected positions for NZE based on Bauer et al. (2007), i.e. relatively close DRESS/TRAP, and a relatively central KIT. As Lorde hails from Takapuna, a coastal suburb that was not included in Ross’s (2018) tests, it is possible that the young speakers in Takapuna have not (yet), or only partially started the sound changes these vowels are engaged in in Auckland.

 However, when changing to singing mode, the data shows that there are considerable differences for both artists. The vast majority of sounds is produced with an increased F1 and F2, which indicates that, when singing, the artists tend to employ a more open upper vocal tract setting, with generally more fronted realizations. This finding is in line with the data found in Gibson (2010). In his study, he argues that a potential reason for this is inherent to (the physical nature of) the singing mode, an argument that was also made in Simpson’s study (1999). This has some basis in vocal technique literature. While there is no evidence of shifting F2 in the literature, it is suggested that F1 could be raised as a result of F0 raising, as this can cause raise of the larynx in non-classically trained singers as well as an increased opening of the jaw. These would cause a shorter vocal tract, and more open vowel settings, respectively (Austin, 2007; Gibson, 2010; Howard, 2009). Data from the Complete Vocal Institute (McGlashan & Sadolin, 2010), who developed the leading singing methodology CVT, support this claim. According to CVI, there are 4 vocal modes in singing, each with their own characteristics and associated vowels. The mode Edge (e.g. Complete Vocal Institute, n.d.), for example, has a high larynx position, and requires that you use more open front vowels, as the use of other vowels this mode could result in vocal damage. However, as Gibson (2010) found that while there is an effect of F0 and duration on F1 in some cases, these cannot be the sole reason for the shifts. Additionally, as the shifts found in this study are not necessarily in the direction the literature would suggest, other factors must be at play here.

As the previously discussed explanations for these shifts are insufficient, they must be evidence for style shifting. For both artists, the fronting of the KIT vowel, the lowering of the DRESS, TRAP, and, quite possibly, the LOT vowels, the backing of GOOSE all indicate a shift from standard NZE towards standard realizations in SSBE or AmE. Further evidence is found in the degree of rhoticity, which is not a feature of NZE, but is a feature of AmE. Though in the majority of occurrences non-prevocalic /r/ was not realized, realization doubles for Kimbra and triples for Lorde in the singing condition compared to the spoken condition.

It appears that there may be multiple pronunciation models at play here with multiple conflicting or co-operating sociolinguistic motivations for the use of these conflicting models. It appears that multiple conflicting or co-operating sociolinguistic motivations are at play here. On the one hand, it is apparent that in both modes, a strong sense of New Zealander identity is retained. In speech, this may be influenced by prestige and result in the employment of features from Cultivated NZE. This could indicate a conflicting social class identity, though that social class is still based in the New Zealander group. On the other hand, the singing mode specifically announces a departure from NZE for multiple vowels. While the divergence seen in this study is consistent with Simpson’s (1999) claim of discourse mode as a motivation for variation to some extent, it seems that features from the American pronunciation model is still present in Kimbra and Lorde’s music. Consequently, when singing, there may be multiple conflicting pronunciation models or these artists could have a relatively unified model as a compromise between, for example, features from AmE, SSBE and NZE. The method of the present study cannot provide an adequate answer for which of these is the case, as this requires an examination of the variation rather than a focus on means. A closer look at Lorde’s “Royals”, however, can reveal underlying sociolinguistic motivations for the(se) model(s).

 Lorde’s debut single “Royals”features R&B and pop influences, and its bare composition features a prominent hip-hop beat. This creates an antithetic juxtaposition of genre elements with the theme of the lyrics. The song denounces the extravagance of pop artists and the lifestyle that is described in hip-hop songs, and “decried the pop industry of which it became a part” (Powers, 2013). It paints the picture of what it is like “growing up in New Zealand immersed in American cultural imperialism” (Perpetua, 2013). The song contains lexical and grammatical features associated with American varieties such as AAVE. Examples are velar fronting (*trippin’, trashin’)*, *don’t*-levelling, and brand names that are associated with affluence in American pop and hip-hop culture (*Grey Goose*, *Cristal*, *Cadillac*). The conflict arises when considering Lorde’s style shift when singing. The American model is present here, and the degree of rhoticity in this song reaches 50%. Through this, it becomes apparent that Trudgill’s appeal to acts of identity alone may not be a sufficient explanation here. In this song, Lorde overtly rejects a desire for identification with of hip hop identity and explicates her membership of the New Zealander group.

This is reminiscent of the situation British Punk-rockers found themselves in in the 1970s, in which there were competing pronunciation models. It seems that there is a complex of motivations in effect when non-Americans sing American music genres. As Gibson (2010) argues, when singing in an American-influenced style, a singer may be affected by responsive referee design. This means that in singing pop music, the singer is responsive to the linguistic norm that exercises its influence through an invisible referee, which will result in an undercurrent of American features in singing these styles. This American norm may, then, be reinforced as a result of genre appropriateness. For example, a hip-hop beat is prominently featured in “Royals” which could induce American features in singing. In contrast, the use of NZE in pop music is speech design initiated by the speaker for a certain invisible, but direct, audience and indicates a deviation from the norm (Gibson, 2010). Simpson’s (1999) idea of linguistic variation motivated by field of discourse could influence the singer to exhibit both more NZE and more AmE features in the present study. While the topic of the song in question is a rejection of extravagant American pop and hip-hop culture, it discusses many examples from this culture explicitly. On one hand, then, the frequent use of these cultural markers can reinforce the linguistic norm, while the subject of dismissal of that culture may inspire deviation from that norm and reinforce the New Zealander identity. Overall, the data from this study indicate that these are in conflict with regards to Lorde’s sung pronunciation, as many sounds remain consistent with NZE, but other features suggest a strong move towards AmE.

In summary, the question this study aimed to answer was: Do Kimbra and Lorde modify their respective English varieties when singing compared to speaking? The data presented here indicated that while these artists spoke an NZE variety, their singing pronunciation exhibited multiple features from, presumably, AmE (and to some extent SSBE). This shift is consistent with the hypothesis that the American pronunciation model would, in some shape or form, still be present in these artists’ sung pronunciation. Multiple sociolinguistic motivations for this phenomenon were explored. For example, the use of these linguistic features as linguistic acts of identity, the influence of linguistic norms as an invisible referee on a speaker or singer, the singing mode as opposed to the speaking mode, genre appropriateness, and the lyrics’ field of discourse.

1. **Conclusion**

The present study has aimed to provide insight into the more recent sociolinguistic situation in New Zealand pop music. This study is not without its shortcomings, however. For example, it is limited to a phonetic analysis of songs by two New Zealand singers. This phonetic analysis had potential issues with interference from backing music and vocals in terms of formant analysis. The present study attempted to counteract this by isolating the vocals, which resulted in a loss of quality of the recording. This could be remedied, perhaps, by obtaining and analyzing the (solo) lead vocal recording used in the song’s mix. Furthermore, rhoticity could have been explored more extensively, and could be indicated more accurately by considering its acoustic correlate, the F3. The present study could, however, be considered a preliminary study of sorts. Further research can build upon this study in several ways.

First, entire albums could be analyzed. This would yield a bigger data set for a more extensive and reliable analysis and make a stronger case for discourse motivated language variation. It could also open up the possibility of observing real-time intraspeaker variation in speech and/or music over time through longitudinal data. Kimbra, for example, moved to Australia before *Vows* and has lived in the United States afterwards. Analyzing her later music could provide evidence for variation in her speech and singing. For instance, this could indicate that she shifts towards the variety that is spoken where she resides, or, conversely, that she emphasizes her New Zealander identity by predominantly using NZE features.

 Second, the present study primarily focused on a phonetic analysis of speech and singing pronunciation. By featuring other aspects of grammar more prominently, a stronger case could be made for variation as a result of genre appropriateness. The previously discussed example of Lorde’s “Royals”*,* with hip-hop beats and AAVE features, presents such an interpretation which may be of value in future research on the same topic.

 Third, the present study was an investigation into a more recent state of NZE in pop music. As mentioned before, Gibson (2010) stated that the American model was still present in New Zealand pop artists, and that a new generation of New Zealand pop stars could change this. The data of this study (and perhaps a future study with a larger scope on the same subject) could then be a catalyst for a paper that reviews the studies done on NZE in pop music over the last decades. This could provide valuable information on both the current state of the American model and how it has developed diachronically in New Zealand pop music.

**References**

Audacity Team (2020). Audacity (version 2.4.2) [Audio editing software]. Retrieved 26 June 2020 from http://www.audacityteam.org

Beal, J. C. (2009). "You're not from New York city, you're from Rotherham": Dialect and identity in British indie music. *Journal of English Linguistics, 37*, 223-240. doi:10.1177/0075424209340014

Bauer, L., & Warren, P. (2004). New Zealand English phonology. In B. Kortmann, W. Schneider, K. Burridge, R. Mesthrie & C. Upton (Eds.), *A handbook of varieties of English: A multimedia reference tool* (pp. 580-602). Berlin, Germany: Mouton de Gruyter.

Bauer, L., Warren, P., Bardsley, D., Kennedy, M., & Major, G. (2007). New Zealand English. *Journal of the International Phonetic Association*, *37*, 97-102. doi:10.1017/S0025100306002830

Bell, A. (1984). Language style as audience design. *Language in society*, *13*(2), 145-204. doi:10.1017/S004740450001037X

Boersma, P., & Weenink, D. (2020). Praat: doing phonetics by computer (Version 6.1.16) [Computer program]. Retrieved 3 February 2020 from http://www.praat.org/

Complete Vocal Institute. (n.d.) *Description and sound of Edge.* CVT research. https://cvtresearch.com/vocal-modes/edge/description-and-sound/

Coupland, N. (2009). The mediated performance of vernaculars*. Journal of English Linguistics, 37*(3), 284-300. doi:10.1177/0075424209341188

Durant, A. (1984). *Conditions of music.* London, England: Macmillan. doi:10.1017/S0261143000006668

Frith, S. (1988). *Music for pleasure*. Cambridge, English: Polity Press. doi:10.1017/S0261143000003639

Gibson, A. M. (2005). Non‐prevocalic /r/ in New Zealand hip hop. *New Zealand English Journal*, *19*, 5-14.

Gibson, A. (2010). *Production and perception of vowels in New Zealand popular music* (Doctoral dissertation, Auckland University of Technology).

Giles H., & Powesland P. (1997) Accommodation theory. In Coupland N., Jaworski A. (Eds.) *Sociolinguistics. Modern linguistics series* (pp. 232-239)*.* London, England: Palgrave. doi:10.1007/978-1-349-25582-5\_19

Gordon, E., & Maclagan, M. (2004). Regional and social differences in New Zealand: phonology. In Schneider, Edgar W., Burridge, K., Kortmann, B., Mesthrie, R. & Upton, C. (Eds.), *A handbook of varieties of English* (pp. 603-613). Berlin: Mouton de Gruyter. doi: 10.1515/9783110175325

Kimbra. (2011). *Vows* [Album]. Auckland, New Zealand: Warner Records.

Kimbramusic. (2011, August 30). *Kimbra talks about her favourite youtube videos* [Video]. YouTube. https://www.youtube.com/watch?v=Bn0RTu9A4sw

Lorde. (2013). *Pure Heroine* [Album]. Auckland, New Zealand: Universal Music.

Lorde Updates. (2013, August 24) *Lorde interview with Polly on ZM Radio New Zealand* [Video]. YouTube. https://www.youtube.com/watch?v=Oy2kf3R\_XTg

McGlashan, J., & Sadolin, C. (August, 2010), *Laryngeal gestures and laryngograph data associated with the four vocal modes as described in the Complete Vocal Technique method of singing teaching.* BVA choice for voice conference, London, U.K.

Newshub. (2015, May 4) *The story of Lorde* [Video]. YouTube. https://www.youtube.com/watch?v=4IsjGOURx20

Nielsen SoundScan, Nielsen BDS. (2013, January 4) The Nielsen company & Billboard's 2012 music industry report. *Business Wire.*

www.businesswire.com/news/home/20130104005149/en/Nielsen-Company-Billboard%E2%80%99s-2012-Music-Industry-Report

O’Hanlon, R. (2006). Australian hip hop: A sociolinguistic investigation. *Australian Journal of Linguistics*, *26*(2), 193-209. doi:10.1080/07268600600885528

Perpetua, M. (2013, November 7) *Lorde’s “Royals” isn’t anti-rap, it’s anti-imperialism.* Buzzfeed. https://www.buzzfeed.com/perpetua/lordes-royals-isnt-anti-rap-its-anti-imperialism

Powers, A. (2013, December 5). *Lorde Sounds Like Teen Spirit.* NPR. https://www.npr.org/sections/bestmusic2013/2013/12/04/248864263/lorde-sounds-like-teen-spirit?t=1597495112239

Ross, B. (2018). *An acoustic analysis of New Zealand English vowels in Auckland.* [Master’s thesis, Victoria University of Wellington]. ResearchArchive – Te Puna Rangahau

Simpson, P. (1999). Language, culture and identity: With (another) look at accents in pop and rock singing. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication*, *18*(4), 343-368. doi:10.1515/mult.1999.18.4.343

Tabouret-Keller, A., & Le Page, R. B. (1985). *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge University Press.

Timt4stic. (2011, May 8). *Kimbra interview on TVNZ 06-05-2011* [Video]*.* YouTube. https://www.youtube.com/watch?v=dNH-2QFEsfY

Trudgill, P. (1983) Acts of conflicting identity: The sociolinguistics of British pop-song pronunciation. In Trudgill, P. (Ed.) *On dialect: Social and geographical perspectives* (pp. 251-265). Oxford, England: Basil Blackwell.

Warren, P. (2018). Quality and quantity in New Zealand English vowel contrasts. *Journal of the International Phonetic Association*, *48*, 305-330. doi:/10.1017/S0025100317000329

Watson, A. (2020, February 13) Global record companies market share 2018. *Statista*. www.statista.com/statistics/422926/record-companies-market-share-worldwide-physical-digital-revenues/

Wells, J. C. (1982). *Accents of English*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511611759

Williams, G. (2011, September 1). *Kimbra: Vows interview 2-9-11 Radio Wammo Show* [Video]. YouTube. https://www.youtube.com/watch?v=l0tEWvKSwj0