Pop Song English as a supralocal norm

A N D Y   G I B S O N

Macquarie University, Australia

A B S T R A C T

An American-influenced singing accent, referred to here as Pop Song English (PSE), is common in popular music throughout (and beyond) the Anglophone world. This article presents an analysis of the sung pronunciation of two variables (BATH and nonprevocalic /r/) that distinguish New Zealand English (NZE) from American Englishes (AmE). The Phonetics of Popular Song (PoPS) corpus includes 154 performers, structured according to country of origin (NZ and the US) and musical genre (pop and hip hop). An auditory analysis was conducted for each variable, distinguishing between the NZE and PSE/AmE variants. Almost all New Zealand performers adopt the PSE variants at least some of the time, with greater adherence to the American model in pop than in hip hop. In the US, region determines hip hop, but not pop, artists' degree of rhoticity. PSE represents a supralocal norm for pop music, while hip hop artists tend to use their 'own accent'. (Pop Song English, singing accent, rap accent, supralocal norm, nonprevocalic /r/, TRAP–BATH split, intentionality, language performance, pop music, hip hop, responsive style, initiative style)*

T H E   S O C I O L I N G U I S T I C S   O F   P O P U L A R   S O N G

American-influenced phonetic styles exist in popular song beyond the geographic borders of the US, and this international variety is referred to here as *Pop Song English* (PSE).[1] This article explores whether commercial pop singers with different spoken dialects (American and New Zealand Englishes) exhibit a similar phonetic style in their recorded performances. Two variables are analysed, both of which differ substantially between American Englishes (loosely referred to here as AmE) and New Zealand English (NZE). Auditory analyses are conducted for words in the BATH lexical set[2] and at sites of potential nonprevocalic /r/. It is expected that irrespective of their country of origin, pop singers will conform to a similar style. As a point of comparison, commercial hip hop artists are also analysed. Regional identity forms a central theme in hip hop culture (Hess 2009; Gilbers, Hoeksema, de Bot, & Lowie 2019), and greater use of regionally specific variants is therefore expected in hip hop than pop (as was found by O'Hanlon 2006 for Australian music). This article has two fundamental aims:

471

(i) To provide benchmark values for performers of pop music from the US in terms of BATH and nonprevocalic /r/, and to quantify the adoption of these PSE features by NZ pop singers.

(ii) To compare the extent of regional variability in hip hop vis-à-vis pop.

Genre as the primary social variable structuring singing accents

Traditionally, singing has provided communities with a way to form social bonds (Watts & Andres Morrissey 2019). The commercialisation of music over the course of the twentieth century, however, has led to global networks of music production and consumption. Through dominance beginning in the early stages of recorded popular music, the US became and remained the centre of commercialised culture extending throughout, and often beyond, the Anglophone world. The American-derived varieties of English used in mass-distributed recordings took root as part of the aesthetic of rhythm & blues, country, jazz, and rock & roll. This project focuses on singing that is commercialised and marketed. Commercial music only comprises a subset of 'song', of course, which ranges widely in function, from lullabies to national anthems. For music created within the 'music industry', genre is a primary structuring force, particularly in the marketing of music to consumers.

Coupland (2011:573) theorised popular song as a 'field of performance organised according to genre', where place is understood as a sociocultural context rather than as a specific region or nation. Rather than focusing on the geographic origins of singers, a dialectology of popular song might be better organised primarily around genre. Genre determines both a range of different accent norms as well as the degree to which a singer's 'own accent' is licensed in song. This article assesses the degree of difference between US and NZ performers in pop (predicted to have a strong supralocal norm, PSE) and hip hop (predicted to have local accent features).

There are many styles of music that either exhibit strong regional variation or have non-US dialect targets (see Westphal & Jansen 2021 for a review). For example, there is an emphasis on the use of regional dialects in the folk song traditions of the British Isles (Watts & Andres Morrissey 2019), while choral singing targets Southern British English features (Wilson 2017), in a context where there is an emphasis on group cohesion in vowel production (Wray 1999). Amongst music genres that are commercial but not 'pop', reggae has its own cultural centre, with artists from outside of Jamaica using phonetic, morphosyntactic, and lexical features of Jamaican Creole and Jamaican English (Gerfer 2018; Westphal 2018). In punk, place and class meanings are foregrounded through a range of semiotic tools (including accent) to demonstrate opposition to normative social structures (Trudgill 1983; Coupland 2011).

Hip hop emphasises both the authentic representation of self and resistance against the mainstream. In hip hop communities around the world, language and dialect mixing represent 'glocal' cultural practice, as artists carve their place in a

transcultural community (Mitchell 2008; Pennycook & Mitchell 2009; Williams 2017; Gilbers et al. 2019). Cutler (2014) has explored questions around authentication for white rappers in depth. Discussing Cutler's work, Pichler & Williams (2016:562) state that while some white rappers 'authenticate by highlighting closeness to African-American street culture, others authenticate by signaling honesty about their own (white, middle-class) background'. While there is diversity in the dialects of English used in popular music, structured primarily according to genre, it is actually the homogeneity of styles which is striking when listening to pop singers from a wide range of geographic origins. The adoption of PSE by non-American singers has been the focus of the majority of sociolinguistic work on singing accents.

The foundational study of American influence on the pronunciation of English in popular music (Trudgill 1983) identified the use of 'Americanisms' in songs by a range of British singers in the 1960s and 1970s. Trudgill found that this American influence appeared to decrease as the 1960s went on, in part due to the massive commercial success of The Beatles, making the UK a cultural centre in its own right. American influence has, however, survived the intervening fifty years of commercial popular music, and remains strong. Beal (2009) and Gibson & Bell (2012) suggest that in the early twenty-first century, the shifts to 'American' features in popular song performances happen largely unconsciously. It is the use of one's 'own' phonetic style in song that takes effort and conscious control.

Pop Song English exists alongside Hip Hop Nation Language (HHNL), which is derived from African American English (AAE) and has become an important part of hip hop culture worldwide (Alim, Ibrahim, & Pennycook 2009). Much of the linguistic research on hip hop focuses on higher domains of language including multilingualism and lexical choices, showing the interplay of the local and the global in situated hip hop practice. In terms of phonetic style, PSE and HHNL share some core aspects of the phonology of AmE, such as not having the TRAP–BATH split (described below), while diverging on others, such as the degree to which they exhibit nonprevocalic /r/.

Vocal artists whose spoken style is phonologically distinct from PSE, and who use their 'own accent' in singing or rap tend to draw attention from fans and the media (and indeed sociolinguists). People tend to notice when an artist 'has an accent'. It is against a landscape of uniformity that such divergences from the PSE norm become marked. Perhaps as a consequence of this markedness, most sociolinguistic research on singing in popular music has focused on a single artist from outside of the US: Beal (2009) and Flanagan (2019) on Alex Turner of Arctic Monkeys; Bekker & Levon (2020) on Die Antwoord; Eberhardt & Freeman (2015) on Iggy Azalea; Jansen & Westphal (2017) on Rihanna; a series of papers by Konert-Panek (2017a,b, 2018) on Amy Winehouse, Adele, and Joe Elliott of Def Leppard; and Duncan (2017) on Keith Urban. Other studies have compared a small number of artists (Trudgill 1983; Simpson 1999; Coddington 2004; Gibson 2005, 2011; Andres Morrissey 2008; Coupland 2011; Gibson & Bell 2012), with the focus generally being on how non-US artists adopt features

of AmE in their singing accent. Few studies have compared a large number of performers from different genres of music (an exception being O'Hanlon 2006) or from different regions of the US (though see Gilbers et al. 2019, showing adherence to local speech styles in rap performances) and few directly compare US and non-US artists. While Gibson & Bell (2012) conducted a controlled comparison of singing and speech, it was lacking a comparison with singers from the US, the presumed 'homeland' of PSE. The lack of US artists in the sociolinguistics of popular music is a gap in the literature that this article seeks to address. Duncan's (2017) study of Keith Urban (an Australian country singer) and three singers from the South of the US covered both key dimensions of comparison: direct comparison of singing and speech within individuals, and direct comparison of US artists and non-US artists, albeit at a small scale.

A clear weakness of this research programme, and one continued by the present study, is the focus on commercial popular music that is performed in English by people who are native speakers of English. A sociolinguistics of popular song needs to cast the net much wider, considering performers who speak English as a second or foreign language (see e.g. Bell 2011; Zhou & Moody 2017; Hermastuti & Isti'anah 2018) and, crucially, sung performance in languages other than English (e.g. Yaeger-Dror 1991, 1993).

Westphal & Jansen (2021) review research into the sociolinguistics of popular music, illustrating both the homogeneity of accents in commercial pop, and the ability of popular music to put a diverse range of local varieties on a global stage. The existing research tends to rely on qualitative analysis of isolated examples. The present article thus aims to fill two gaps in the literature, providing a quantitative description of two of the USA–5 variables (Simpson 1999) in the performances of US artists, as well as comparing their performances to artists from New Zealand selected using the same protocols. The present analysis is still limited however, since it does not include a comparison to the speech of the artists analysed. I turn now to a description of the study presented in Gibson & Bell (2012), where a direct comparison of singing and speech was conducted.

The question of intention: PSE as a default style

Jansen (2018) explored British listeners' attitudes to singing accents, and concluded that an Americanised accent is the default, expected style in popular song, despite some positive appraisal of accents that diverged from the norm. An important theoretical construct useful to the relationship between language use and intentionality is Bell's (2001) distinction between the *responsive* and *initiative* dimensions of style. A responsive style shift is one which is appropriate and predictable given the interlocutors and the context, while an initiative style shift is one which changes the communicative context in some way or reframes the interlocutors' identities or roles. Gibson & Bell (2012) argued that the use of PSE in song is actually a responsive style, even if it involves shifting away from one's spoken

style, because of its predictability in, and appropriateness to, the pop song context. Using a regionally marked variant, by contrast, is deemed an act of initiative style-shifting, even though it involves the use of a feature consistent with a performer's own regular speech style. In the remainder of this section, I review Gibson & Bell (2012) in some detail, considering the question of INTENTION in the adoption of PSE by singers whose spoken dialect differs to PSE.

Gibson & Bell (2012) showed that New Zealand singers adapt their entire vowel space when singing, rather than adopting only salient 'Americanisms'. By conducting an acoustic analysis of the singing and speech of three NZ singer-songwriters, and by interviewing them about their attitudes and experiences, Gibson & Bell (2012) argued that the 'default' singing accent for these New Zealanders was derived from AmE. Gibson & Bell (2012) included some variables that belong to the USA–5 (LOT and PRICE) along with six other vowels that are less likely to attract stereotype levels of awareness as they relate to NZE and AmE (DRESS, TRAP, THOUGHT, START, GOOSE, and GOAT). Acoustic analysis revealed a dramatic style-shift between speech and singing across all variables. Figure 1, reproduced from Gibson (2010), shows these differences for one of the singers, Dylan Storey.

The differences between singing and speech are dramatic, and this is in part due to factors relating to singing technique. Importantly, there is a tendency for greater sonority in song (Andres Morrissey 2008; Gibson 2010), including greater jaw opening, resulting in more open vowels and higher F1 values. There may also be an overall raising of formant values due to higher fundamental frequencies in singing, and thus higher harmonics at which formants can be amplified. Not all of the differences between singing and speech in Figure 1 can be explained by singing technique, however. Some differences are clearly dialectal. There is an opposing direction of F2 movement in the trajectories of the GOAT and GOOSE[3] vowels, for example, that reflect differences between NZE and AmE/PSE.

The three singers were interviewed to examine their intentions around identity projection. One of the singers said he had not thought much about his singing accent and had no desire to sound like a New Zealander in song, while the other two singers both stated that they would like to use NZE in their songs but found it difficult to do so. Despite these differing identity orientations, the vowel realisations produced in song by the three singers were strikingly consistent. Of the two singers who stated having some desire to use NZE in their singing, both produced occasional NZE vowels in song and reported conscious awareness of producing those vowels with a NZ accent, for example through re-recording a particular vocal part line by line to achieve a sung NZE style. These counter-examples to the PSE default showed that while these singers are CAPABLE of using NZE in song, doing so requires effort and awareness.

The conclusions of Gibson & Bell (2012) can be summarised as follows. A levelled variety of American-derived English, which I refer to here as Pop Song English, constitutes a supralocal norm for singing in popular music (with exceptions according to musical genre, however). This variety is the default singing style for NZ singers,
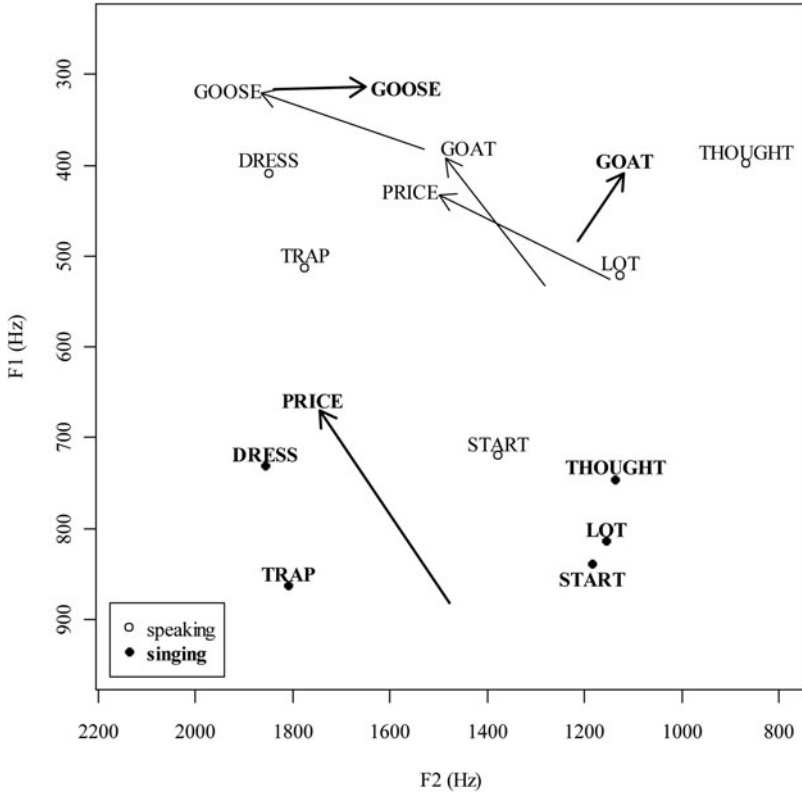
FIGURE 1. Mean F1 and F2 of sung (*n* = 116) and spoken (*n* = 161) vowels for Dylan Storey, reproduced from Gibson (2010). Labels for diphthongs at arrow heads.

affecting the entire vowel space, rather than being a stylisation restricted to prototypical Americanisms. The use of PSE is thus theorised as a responsive style (Bell 1984), as the least marked phonetic style in the context of popular song.

Use of NZE phonetic variants in singing is done intentionally, for example, to project an 'authentic' identity. It is an initiative stylistic move and a case of referee design (Bell 1984) for which the referee is the performer's own spoken style. 'Own-accent' singing thus represents an initiative style-shift, with an implication of heightened intentionality. As such, the use of 'own-accent' features is more likely to happen on more sociolinguistically salient variables, or in more cognitively salient environments (cf. Yaeger-Dror 1991, 1993).

SOCIOLINGUISTIC VARIABLES FOR ANALYSIS

The variables to be studied in this article are BATH and nonprevocalic /r/. These are both members of the group of variables studied by Trudgill (1983) and

476   *Language in Society* **53**:3 (2024)

subsequently labelled the USA–5 by Simpson (1999). These variables (along with intervocalic /t/ flapping, unrounded LOT, and PRICE monophthongisation, not addressed here) were selected by Trudgill because they were deemed to be salient markers of the distinction between British and American English dialects. Trudgill's (1983) study suggested that British pop and rock artists were intentionally imitating American performers in their adoption of these 'Americanisms'. As a mannered act of identity (Le Page & Tabouret-Keller 1985), this imitation was subject to limitations, evidenced for example by cases of hyper-correct /r/-insertion by The Beatles and Cliff Richards. Such cases of phonetic overshoot provide evidence of a performer's intention to target a dialect (Agha 2005; Bell & Gibson 2011; Gibson 2011), and so they provide good evidence that, at least in the 1960s, British artists were 'trying to sound American'. I would expect (though this remains an empirical question for future research) that such cases of overshoot will have decreased steadily over time as succeeding generations of singers have become more NATIVE-LIKE in PSE, having spent their critical period of language acquisition exposed to a relatively consistent model of English in the popular songs they hear around them growing up.

BATH

The first variable under analysis in this article involves words such as *can't*, *dance*, *past*, and *laugh*. In dialects such as Standard Southern British English and NZE, these words are realised with a long open vowel, rhyming with words like *heart* and *calm*. In North American dialects, they are realised with a short front vowel (to brush aside the complex allophonic and lexical conditioning of TRAP), and rhyme with words such as *hand* and *cap*. Realisation of words in the BATH lexical set[4] with [æ] has frequently been discussed in studies of singing accents, as one of the USA–5 features adopted by singers outside of the US. O'Hanlon (2006), for example, found that in Australian popular music, 100% of BATH tokens were realised with TRAP (the PSE variant) in pop compared with only 11% in hip hop. Coddington (2004) found that 56% of the BATH tokens analysed were realised with /æ/ in a sample that included pop, rock, and punk artists from New Zealand. When interviewed about their singing accents, five of the eight artists mentioned awareness of the BATH variable, suggesting a high level of salience for this variable amongst NZ performers.

BATH represents something of a special case for the analysis of singing accents due to the presence of the TRAP–BATH split in NZE and its absence in AmE (for a description of the process leading to this outcome, see Wells 1982). There is a cross-dialectal difference at the phonemic level, with BATH words aligning with PALM (realised as /aː/ ~ /ɑː/) in NZE and TRAP (realised as /æ/) in AmE.[5] Given this difference of phonemic alignment, the variant of BATH chosen affects the rhymes that an artist can or can't use, and is therefore particularly likely to gain a New Zealand singer or rapper's attention during the process of writing lyrics. While gradient acoustic

variation no doubt exists, the choice between variants is likely to be relatively categorical. Performers may have particularly high levels of awareness of the variation of BATH between AmE and NZE for multiple reasons. Listeners are particularly sensitive to variability that crosses phoneme boundaries (Liberman, Harris, Hoffman, & Griffith 1957) and tend to minimise the perception of differences within phoneme categories (e.g. Best 1994). Another reason for potentially heightened awareness comes from the uniformity of both American and NZ Englishes in their realisation of this variable (categorical alignment with TRAP for BATH words in AmE and categorical alignment with PALM for BATH words in NZE). One of Le Page & Tabouret-Keller's (1985) riders to linguistic modification is the ability to understand the model, and for BATH, the model distinguishing AmE from NZE is simple and consistent.

Nonprevocalic /r/

Like BATH, rhoticity (that is, the production of /r/ in nonprevocalic environments) may be relatively cognitively accessible to performers. Presence or absence of nonprevocalic /r/ has stereotype status in distinguishing North American dialects from Southern British English and Southern Hemisphere Englishes. New Zealand English is largely non-rhotic, except for a small population in the south of the South Island (Villarreal, Clark, Hay, & Watson 2021), and partial rhoticity in Pasifika communities (Gibson 2016; Marsden 2017), particularly in the NURSE lexical set. The US, by contrast, is largely rhotic, with exceptions in New England and New York (Becker 2014), the South (Thomas 2003; Carmichael 2017), and in AAE (Wolfram & Thomas 2002).

Adoption of partial rhoticity by non-rhotic singers is another of the USA–5 features. It was included in O'Hanlon's (2006) study of Australian music, where hip hop artists barely used any nonprevocalic /r/ (2%), pop-rock, alternative, and punk performers used somewhat more /r/ (10%) and pop singers used the most (24%). Coddington's study of NZ pop, rock, and punk artists found that only 4% of tokens had a clearly pronounced nonprevocalic /r/, with a further 4% of tokens having a 'slightly audible hint of /r/' (2004:60). For the one artist whose genre was described as commercial pop, the rate was 15% (plus 6% slightly audible /r/). A study of NZ Pasifika hip hop artists (Gibson 2005) showed that NURSE words were consistently rhotic, while all other vowel environments were /r/-less.

The existence of variation in different parts of the US means there is scope for testing the relationships between musical genre and the speech styles of performers' communities. PSE was at least to some extent derived from (non-rhotic) African American and Southern varieties of American English. These origins may have led to lower rates of nonprevocalic /r/ in PSE today than in rhotic varieties of AmE. Given its roots in African American culture, lower rates of rhoticity are also expected in hip hop than pop. Since rhoticity has clear regional variation within the US, the interaction of genre with artists' region of origin is examined for this variable amongst the US artists in the corpus, in addition to the comparison of the US with NZ.

RESEARCH QUESTIONS

A sociolinguistics of popular song has many big questions to explore, including not only the phonetic consequences of singing itself, but also the tensions between genre and geography, between learned routines and intentional innovation, and between adherence to genre-based norms and the expression of the autobiographic voice. The present article aims to provide a stepping-stone to these larger issues by examining a carefully selected sample of songs to explore three specific research questions.

(i) Do NZ pop singers produce the PSE variants of BATH and nonprevocalic /r/ at similar rates to US pop singers?
(ii) Do NZ hip hop artists produce the PSE variants of BATH and nonprevocalic /r/ at lower rates than NZ pop singers?
(iii) With respect to nonprevocalic /r/, do US hip hop artists adopt a level of rhoticity that reflects their place of origin?

INTRODUCING THE POPS CORPUS

The Phonetics of Popular Song (PoPS) corpus in its current form is made up of 190 vocal performances by 154 artists. It is structured by genre (pop and hip hop), country of origin (NZ and the US), ethnicity (Pākehā and Māori/Pasifika in NZ, and European American and African American in the US) and gender (male and female in pop, but only male in hip hop since very few female hip hop tracks were revealed with the song selection methods described below). The number of songs and artists in each of these demographic cells is summarised in Table 1.

Methods of song selection

Avoidance of selection bias was one of the primary motivations in developing the methodology for song selection, which proceeded systematically using the NZ singles charts maintained by Recorded Music New Zealand[6] with the majority of songs coming from 2015–2017. Setting up in advance a stringently defined set of rules to govern the selection of songs, I made myself as 'tasteless' (Brooks 1982) as possible. That is, I did not allow my own judgements about the worthiness of a given song for study to guide selection decisions. Since the primary interest of this project was to focus on the music to which New Zealand listeners are exposed, these charts were used to find the songs by both the US and NZ artists, using the following inclusion criteria.

- COUNTRY OF ORIGIN: Artist must have grown up in NZ or the US. There is debate about the critical/sensitive periods for language and dialect acquisition (Werker & Hensch 2015). To be included in the corpus, each performer had to have moved to NZ/US by the age of five.

TABLE 1. *Number of songs in each cell of the PoPS corpus, with number of unique artists in brackets.*

Country of origin	Ethnicity	Female pop	Male pop	Male hip hop	Total
NZ	Māori/Pasifika	20 (13)	17 (12)	19 (17)	56 (42)
NZ	Pākehā	15 (13)	16 (11)	13 (10)	44 (34)
US	African American	15 (11)	15 (10)	15 (15)	45 (36)
US	European American	15 (15)	15 (15)	15 (12)	45 (42)
Total		65 (51)	63 (48)	62 (54)	190 (154)

- GENRE: The genre of the artist had to be either pop or hip hop/rap on the artist's page in Apple Music. The decision to use Apple Music genre was made for replicability and simplicity, since Apple Music is rare amongst online music platforms in allowing only one genre label per artist.[7]
- ETHNICITY: Artists were placed into one of four broadly construed ethnic groups: NZ Māori or NZ Pasifika, NZ Pākehā (New Zealanders of European descent), African American, and Americans of European descent. For an analysis of results with respect to artist ethnicity, see Gibson (2020).
- GENDER was treated as binary, and I acknowledge that this binary categorisation is reductive and problematic.

Other inclusion criteria were predefined to clarify the number of tracks that could be included from a given artist and how to deal with tracks that have multiple vocal performers.

- REGION WITHIN THE US: While US artists were not selected in order to cover a certain range of regional backgrounds, it was decided that for the analysis of nonprevocalic /r/ this information needed to be ascertained. A binary distinction was made between more and less rhotic regions of the US, grouping performers from West Coast states and from areas in the Midwest (including towns as far east as Pittsburgh, Pennsylvania) in one category, and those from the East (including towns in eastern Pennsylvania) and the South in the other category. Artists who moved between regions during childhood were removed from the analysis of regional differences amongst the US artists.

Songs identified for inclusion were purchased through Apple Music, converted to wav files and imported into Praat (Boersma & Weenink 2019). Lyrics were transcribed and manually time-aligned to the soundfile at roughly one-line intervals, with identically repeated sections excluded from analysis. Audio files and Praat textgrids were uploaded to LaBB-CAT (Language Brain and Behaviour Corpus Analysis Tool, Fromont & Hay 2012), where the corpus is stored and managed. The transcripts were force aligned at the phoneme level using HTK (Hidden Markov Model Toolkit). Despite the fact that the vocals appear in the context of instrumentation, HTK alignment was impressively accurate, making it easy to search for and precisely locate variables of interest.

ANALYSIS OF THE POPS CORPUS: METHODS

BATH

An auditory analysis was carried out for the 301 tokens of BATH that occurred in the corpus. The initial aim was to designate each token as having either the phoneme /æ/ (that is, cases where BATH words align with the TRAP lexical set) or /ɑː/ (where BATH words align with the PALM lexical set). However, three categories were needed to capture the variation, with nineteen tokens being realised as an up-gliding diphthong, rather than aligning with TRAP or PALM. All of these tokens occurred in the word *can't*. For the binary analysis, these diphthongal tokens were included in the TRAP category. In this analysis of BATH (and also for nonprevocalic /r/, below), function words are included in the datasets. Care was taken to exclude items realised as unstressed and having a reduced vowel. Vowel reduction may be rarer in song than in speech, where each syllable has a rhythmic function. Given the limited size of the lexicon in pop songs (Murphey 1992), function words are deemed to be an important part of the dataset, and any systematic variation that they exhibit will be controlled for with the inclusion of a random intercept for word in statistical models, wherever this does not lead to convergence issues.

Nonprevocalic /r/

An auditory analysis was conducted for 3,659 tokens, along with visual inspection of the spectrograms in Praat. Of the 3,659 tokens originally exported from LaBB-CAT, fifty-eight were excluded due to the candidate token being followed by another /r/, or due to mistranscription. A further 359 tokens at sites for potential linking /r/ were also removed from the present analysis, all of which were assessed auditorily to ensure the /r/ was directly followed by a vowel. If there was a pause or prosodic boundary before the following vowel, the token was included as part of the present analysis of nonprevocalic /r/. The results for linking /r/ can be found in Gibson (2020).

Care was taken to provide a quality categorisation of the data into /r/ and /r/-less tokens. In recognition of the fact that /r/ is not a binary variable, but rather a very complex package of both temporal and spectral cues, detailed information was recorded for each token, even though this was ultimately collapsed into a binary /r/ present vs. absent distinction. For the 3,242 tokens, six codes were used to denote the type of realisation. These included one code to mark complete absence of /r/ ($n = 1976$), and three to capture varying degrees of post-vocalic /r/ presence, reflecting the perceived degree of constriction and length of the /r/ (subtle /r/, $n = 156$; moderate /r/, $n = 214$; strong /r/, $n = 539$). In addition to these main categories, there were 324 tokens of rhoticised vowels [ɚ], where more than half of the length of the vowel was perceived to be /r/-coloured. Many of these tokens did not have a post-vocalic consonantal /r/ segment, despite still clearly counting as examples of rhoticity. Finally, there were thirty-three tokens where a vocalic

offglide gave me the initial impression of an /r/ segment, despite the absence of any actual rhoticity. For example, a non-rhotic FORCE vowel realised as [fɔːəs] can be initially misperceived by a non-rhotic listener as containing /r/ if care is not taken.

Ultimately, these six categories were collapsed into a binary analysis. The three categories denoting some degree of consonantal post-vocalic /r/ were grouped with the rhoticised vowel tokens, yielding 1,233 instances of /r/-presence. The non-rhotic offglide tokens were grouped with the no-/r/ tokens, yielding 2,009 /r/-absent tokens.

In addition to the six categories, tokens were flagged in cases where my confidence in the code assigned was low. Across the full dataset (including linking /r/ environments), a total of 538 tokens were marked as uncertain. A further seventy tokens were noted to be difficult to assess due to being obscured by the instrumentation of the song. All tokens marked with one of these flags was subjected to a blind reanalysis, along with a random sample of 150 non-problematic tokens. For this re-analysis phase, a binary assessment of /r/ presence vs. absence was made. For the 150 non-problematic tokens, the check–recheck agreement rate of the two analyses was 97%. For the tokens marked as problematic, however, this re-analysis yielded a lower intra-rater agreement rate of 74%. A third blind listen was conducted for those tokens where the first two analyses differed, and the majority code was then entered as final. Any tokens that were marked as being obscured by the instrumentation on both the first and second pass were excluded from the dataset ($n = 16$).

Statistical analysis methods

For both the BATH and nonprevocalic /r/ analyses, binomial generalised linear mixed effects regression models are fit with the *lme4* package in R (Bates, Mächler, Bolker, & Walker 2015; R Core Team 2019). For BATH, the dependent variable is the likelihood of realising a BATH word with /æ/ (the TRAP variant). For the rhoticity models, the dependent variable is the likelihood of /r/-presence. None of the statistical models presented in the results section below should be construed as confirmatory hypothesis testing, but rather as exploratory analysis of the corpus data (for a discussion of the distinction between exploratory and confirmatory data analysis see Nosek, Ebersole, DeHaven, & Mellor 2018). During data exploration, multiple models were run on various subsets of data, so all *p*-values should be considered anti-conservative. Additionally, most models are fit with only random intercepts and not with slopes and are thus also anti-conservative for this reason. Future research, however, can determine testable hypotheses on the basis of these results.

To explore the first two research questions, the full datasets for BATH and nonprevocalic /r/ are each tested in a model that includes the interaction of genre with country of origin. Genre is a factor with two levels: hip hop (the reference level)

482   *Language in Society* **53**:3 (2024)

and pop. The singer's country of origin is also a binary predictor, distinguishing NZ (the reference level) and the US. To explore the third research question, regarding regional variation in the US with respect to rhoticity, a model is fit on a subset of data that includes only those sixty-five US artists for whom reliable information about region of origin could be obtained. In this model, the interaction between genre and region of origin is tested. Region of origin is a two-level factor distinguishing rhotic parts of the US (the reference level, labelled West/Midwest) from less rhotic areas (labelled South and East). The only linguistic-internal constraint that was deemed to be critical for inclusion in any of the models was the vowel environment for potential cases of nonprevocalic /r/. Since the NURSE environment strongly favours rhoticity, a binary distinction for vowel environment is included in the rhoticity models. This is a two-level factor distinguishing between tokens that occur in the NURSE lexical set from those that occur in any other environment (the reference level).

Random intercepts for performer and word are included in all models, unless their inclusion leads to non-convergence. The intercept for word groups all words that only occurred once into a single level. This way, the intercepts on word account for idiosyncratic behaviour in words that occur multiple times in the dataset, but are not overly sensitive to the peculiarities of words that appear only once. In the rhoticity model including all data, a slope for NURSE on performer is included, given potential differences in the degree to which NURSE favours rhoticity across individuals. For the model exploring regional variation in rhoticity amongst US artists, however, the slope for NURSE could not be included due to non-convergence.

The significance of the genre by place of origin interaction can reveal whether place-based differences are greater in one genre than the other. To assess these interactions in more detail, pairwise comparisons are run on each model (using the *emmeans* package, Lenth 2020) to provide an indication of the significance of differences between groups (bearing in mind once again that this is in the context of an exploratory, not a confirmatory, analysis).

ANALYSIS OF THE POPS CORPUS: RESULTS

BATH

Across the corpus of NZ and US pop and hip hop performers, 301 tokens of BATH were designated as being realised with either TRAP (/æ/) or PALM (/ɑː/). In these broad terms, 254 tokens (84%) of the BATH words were aligned with the TRAP lexical set (and realised with /æ/), and forty-seven tokens were aligned with PALM, and realised with /ɑː/. Table 2 shows the percentage of tokens realised with TRAP for each combination of genre and country of origin. In the US data, the results are near categorical, with all but three of the 167 tokens realised with /æ/. In NZ songs, the realisation of BATH words with /æ/ is also prevalent, with 67% of the 134 tokens using this PSE variant, though this rate varies according

TABLE 2. *Mean rate of realisation of* BATH *words with* TRAP *(/æ/) for each combination of genre and country, with token counts. Means of by-performer means are also given since token counts vary between performers.*

Genre	Country	Mean of performer means (% TRAP)	N performers	Grand mean (% TRAP)	N tokens
Pop	US	97.1	34	99.0	100
Pop	NZ	78.3	33	73.9	88
Hip hop	US	98.0	20	97.0	67
Hip hop	NZ	48.3	20	54.3	46
Totals			107		301

to genre. Taking the mean of performer means in each genre, the average rate of realising BATH with /æ/ in NZ is 78% in pop, and 48% in hip hop.

Table 2 (along with other tables describing raw results) includes both grand means and means of by-performer means. The means of means are given to reduce the effect of widely varying token counts for different performers. To illustrate the difference, consider the results for the percentage of BATH tokens realised with TRAP in NZ hip hop. The grand mean across all tokens is based on twenty-five out of forty-six tokens (54.3%) having TRAP. Of the twenty artists contributing to this statistic, seven artists only have a single token, while two artists have six tokens each, and thus contribute more to the grand mean than the artists with only one token. Both of those artists with six tokens happen to use TRAP consistently, and thus the grand mean is inflated. The mean of by-performer means is a lower value (48.3%), since the TRAP-using artists with a high token count contribute only once each to the statistic.

The regression model for BATH included a significant interaction of genre with country of origin and a random intercept for performer (see Appendix A for the model summary). Figure 2 shows the fitted interaction from the model, along with a summary of the raw data. Lines drawn between the model predictions (on this and all other figures) for the two genres are included solely to aid visual comparison, not to imply a continuous relationship between the genre categories. The large points (connected by lines) show the model fit, back-transformed from log odds to probabilities. The small points show the mean rate of realising BATH words with TRAP for each individual performer (plotted using the *geom_jitter* function within the *ggplot2* package (Wickham 2016) to spread the points and aid readability). Due to the bimodal nature of the results, with most performers being consistent in their choice of variant, the model makes polarised predictions, near zero and one. The model predicts that US artists, and also NZ pop artists, will realise BATH words with TRAP. NZ rappers, however, are predicted by the model to realise BATH with PALM. Inclusion of the raw data shows that the variation is somewhat more nuanced, with a few NZ pop singers using PALM and several NZ hip hop artists using TRAP, along with six New Zealanders that use both variants. The
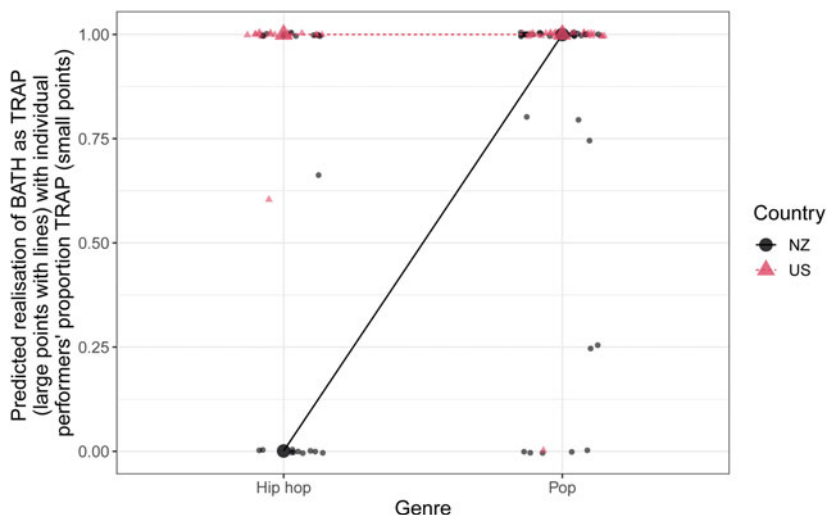
FIGURE 2. BATH model ($n = 301$): Predicted probability of realising BATH with TRAP ($/æ/$) according to genre and country of artist. Lines connect the predictions from the model fit for each genre category, back-transformed to probabilities. Small points (plotted with jitter for readability) show each individual performer's mean rate of TRAP.

pairwise comparison shows no significant difference between NZ and US pop ($p = 0.858$), and a significantly lower likelihood of using TRAP for NZ hip hop as compared to NZ and US pop (both $p < 0.001$), and US hip hop ($p = 0.002$).

Nonprevocalic /r/: Analysis of country of origin across all data

The first of two models exploring nonprevocalic /r/ looks at variation across the entire dataset, comparing NZ and US performers of pop and hip hop. Across the full dataset, there were 1,206 (37%) /r/-ful tokens and 2,036 (63%) /r/-less tokens. The mean rate of /r/ realisation and the number of tokens for each combination of genre and country are shown in Table 3, along with aggregate information for the distinction between NURSE and non-NURSE environments. In the NZ data as a whole, a grand mean of 30% of all tokens were realised with /r/, compared to 45% of all tokens in the US data. The lower value is driven mainly by NZ hip hop artists, with 21% rhoticity, though NZ pop artists also use lower rates of nonprevocalic /r/ (35%) than US artists in either pop (41%) or hip hop (51%). Once again, the grand means are affected by differing token counts for each artist. Looking at by-speaker means reveals a similar rate of 43% rhoticity for both US pop and hip hop. As expected, nonprevocalic /r/ is much more likely to be realised in words in the NURSE lexical set (grand mean 81%) than in other environments (grand mean 27%).

The generalised linear mixed effects model for the likelihood of realising nonprevocalic /r/ included a significant interaction of country of origin with musical

TABLE 3. *Mean % /r/ realisation and token counts for rhoticity data, grouped first according to genre and country, and then according to whether the potential /r/ occurs in a* NURSE *environment or not. Means of by-performer means are also given since token counts vary between performers.*

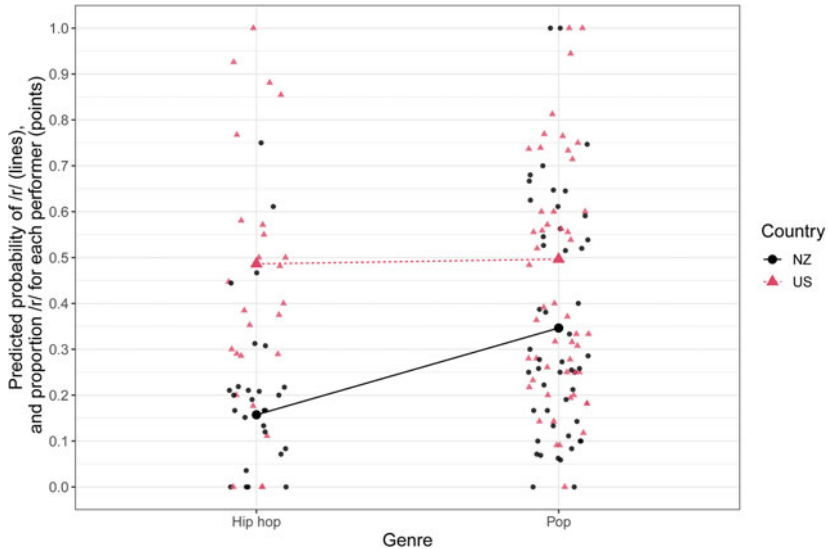| Genre	Country	Mean of performer means (% rhotic)	*N* performers	Grand mean (% rhotic)	*N* tokens
Pop	US	43.3	51	41.0	1021
Pop	NZ	35.2	49	35.7	1016
Hip hop	US	43.2	26	50.9	583
Hip hop	NZ	20.9	27	20.4	622
Totals		36.7	153	37.2	3242
	Lexical set	Mean of performer means (% rhotic)	*N* performers	Grand mean (% rhotic)	*N* tokens
	NURSE	76.9	131	80.5	596
	Other vowels	27.4	152	27.4	2646
	Totals	36.7	153	37.2	3242

FIGURE 3. Rhoticity model for all data ($n = 3242$): Predictions from interaction of genre and country (larger points connected by lines) plotted with individual performers' proportion of /r/-presence (small points).

genre ($p = 0.048$). The favouring effect of the NURSE environment was also highly significant ($p < 0.001$). The model also included random intercepts for performer and word, with a slope for NURSE on performer (see Appendix B for the model summary). The predicted rate of /r/ realisation from the interaction of genre with country is shown in Figure 3, along with the mean rate of /r/ for each performer. Most New Zealand performers produce nonprevocalic /r/ at least some of the time, and rates of /r/ are higher in pop than hip hop. The pairwise comparison shows no significant difference between NZ and US pop ($p = 0.134$), and a significantly lower likelihood of using nonprevocalic /r/ for NZ hip hop as compared to NZ pop ($p = 0.022$), US pop ($p < 0.001$), and US hip hop ($p < 0.001$).

Nonprevocalic /r/: Analysis of regional variation in the US

The second model for nonprevocalic /r/ looks at variation amongst the US artists, considering whether hip hop artists display their regional dialect through rhoticity. The mean rate of rhoticity and number of tokens for each genre by region group is shown in Table 4. The grand mean rates of rhoticity are very similar for artists from the more rhotic (40%) and less rhotic (39%) regions in the context of pop songs, but for hip hop, there is a much lower rate of rhoticity in the non-rhotic regions of the South and the East Coast (26%). Rappers from rhotic regions have a much higher rate of rhoticity (60%) in their rap than pop singers from either region.

TABLE 4. *Mean % /r/ realisation and token counts for rhoticity data from US artists only, grouped according to genre and the performer's region of origin. Means of by-performer means are also given since token counts vary between performers.*

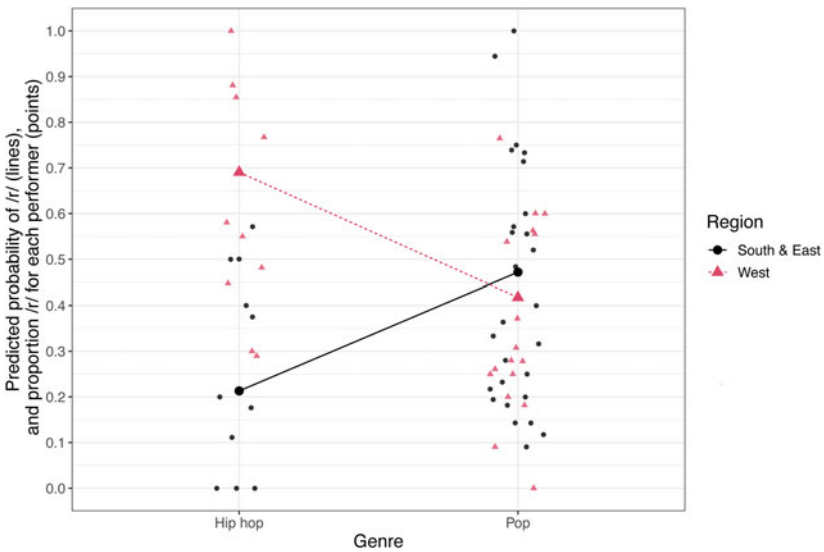| Genre	US region	Mean of performer means (% rhotic)	N performers	Grand mean (% rhotic)	N tokens
Pop	West/Midwest	35.8	17	39.7	312
Pop	South/East	43.1	27	39.2	567
Hip hop	West/Midwest	61.5	10	59.7	365
Hip hop	South/East	25.8	11	25.9	116
Totals		41.1	65	43.7	1360

FIGURE 4. Rhoticity model for US data only ($n = 1360$): Interaction of genre with region (lines) plotted with individual performers' proportions of /r/-presence (points).

The model for the US data included a significant interaction of genre with region ($p = 0.003$), a significant main effect for whether the /r/ was in a NURSE word or not ($p < 0.001$), and random intercepts for performer and word (see Appendix C for the model summary). Figure 4 shows the interaction of genre and region for these US artists. Predicted values are plotted along with the mean rate of /r/ for each participant. The pairwise comparison shows no significant difference between West/Midwest and South/East in pop ($p = 0.956$), but significantly lower likelihood of using nonprevocalic /r/ for South/East hip hop as compared

to West/Midwest hip hop ($p = 0.006$). None of the other pairwise comparisons reached significance.

ANALYSIS OF THE POPS CORPUS: DISCUSSION

BATH

As expected, BATH aligns with TRAP (/æ/) for US artists in both genres, reflecting American Englishes and thus also Pop Song English. Realisation of BATH with /æ/ was also prevalent in the performances by New Zealanders, and this was especially the case in pop. NZ pop singers use less TRAP than US pop singers in terms of raw values, but this difference was not statistically significant. NZ hip hop artists, by contrast used much lower rates of the PSE variant. Some artists adopt the HHNL/PSE variant (the American model is the same in both genres) while others use the NZE variant, possibly as an act of authentication, displaying their 'real' self by using their 'own accent' in their performances.

While NZ pop is strongly influenced by the PSE model, it is not indistinguishable from it. There are several NZ pop artists who do not follow the PSE model. For NZ artists attempting to use their own accent in performance, BATH may be the easiest variable with which to enact this identity goal, because of its likely high level of salience. While the present analysis has not directly probed awareness, my impression is that BATH is a variable where many NZ artists feel they have to make a conscious choice between two highly contrastive, and socio-indexically meaningful, variants.

Nonprevocalic /r/

The results show the adoption of nonprevocalic /r/ by NZ pop vocalists, approaching the PSE norm set by the US pop artists. Hip hop artists, by contrast, appear to have different targets that correspond to their local spoken dialect. The first model showed that NZ hip hop artists had much lower rates of rhoticity than any of the other groups, including rappers from the US. The second model revealed, however, the danger of grouping hip hop artists from both rhotic and non-rhotic regions of the US into a single category. With their region taken into account, rappers from the South and East of the US have a similarly low rate of rhoticity to NZ rappers. Hip hop artists from rhotic areas of the US, by contrast, have higher rates of nonprevocalic /r/ than any other group. Comparing these rappers to the pop singers provides some support for the idea that PSE is less rhotic than would be expected for rhotic varieties of spoken American English. This might reflect the strong influence of both Southern and African American artists in the formation of Pop Song English, and/or it may relate to singing-technique factors such as a preference for sonority.

While the lower rate of nonprevocalic /r/ in NZ than US pop did not reach significance, a more highly powered study would likely find a difference. There are at

least two possible reasons why NZ pop singers have lower rates of rhoticity than US pop singers: first, there could be some degree of intentional own-accent singing; second, there may be imperfect application of the model. The former account suggests that PSE is default, and that in the absence of the intention by some singers to sound like a New Zealander, the rates of rhoticity would be higher. The second account suggests that the NZ pop singers are in fact TRYING to sound like American pop singers, and failing to do so accurately. These options can only be disentangled by finding out about singers' intentions, which is beyond the scope of the present study.

GENERAL DISCUSSION

Taken together, the results of the corpus analysis provide three main findings, relating to the three research questions proposed earlier.

(i) NZ pop singers produce the Pop Song English variants of BATH and nonprevocalic /r/ at rates comparable to US pop singers. BATH is realised with TRAP and partial rhoticity is adopted. Both occur at slightly lower rates in NZ pop than US pop, though these differences were not significant in pairwise analyses of the models.

(ii) NZ hip hop artists produce the PSE variants of BATH and nonprevocalic /r/ at significantly lower rates than NZ pop singers.

(iii) With respect to nonprevocalic /r/, US hip hop artists adopt a level of rhoticity that reflects their place of origin, and both of these rates of rhoticity differ from the PSE norm. Rappers from non-rhotic areas use less /r/ than is used in pop, while rappers from rhotic areas use more /r/ than is used in pop.

This article provides one of the first attempts at a quantitative description of how PSE looks in its 'homeland', for two of the variables most often studied in the sociolinguistics of popular song. For artists from the US, BATH is realised almost categorically as TRAP, irrespective of the musical genre. Rhoticity, by contrast, is much more variable, as indeed it is across varieties of American English. Most sociolinguistic studies of singing accents in popular music have focused on non-US artists, and have sometimes assumed that the PSE model has very high, or even categorical, levels of rhoticity. Such assumptions have been based on a lack of information about the PSE model. Consider O'Hanlon's (2006:200) comment that an Australian singer with 28% rhoticity was 'unable to fully rhoticise her singing' due to a lack of control over production of the variable. In the PoPS corpus, twenty out of the fifty-one US pop artists have a mean rhoticity rate of less than 28%. In light of this finding, it seems less clear that O'Hanlon's Australian singer was unable to accurately emulate the PSE model. It may indeed have been quite a typical performance of PSE.

NZ pop was found to have a slightly lower rate of rhoticity than US pop. This could be taken as a sign that singers are unable to accurately adopt the model, or it could be a sign that some singers are actively shunning Americanisms. In

order to assess questions of intention, a range of variables need to be studied, and their relative degree of salience assessed empirically. Such salience may vary for different lexical items, and according to the context in which a given token occurs (see Yaeger-Dror 1993), in addition to broader variability in salience from one variable to another. Understanding the relative levels of awareness and control performers might have over different variables would aid in the interpretation of their sung performances. If singers are 'trying to sound American', then we would expect greater awareness and control to lead to more successful imitation of PSE, while less salient variables would be produced with NZE. By contrast, if PSE is a default style, and awareness and intention are required in order to 'use their own accent', then we would expect more use of NZE variants on salient variables, and a closer match with the PSE model on less salient variables. BATH and rhoticity (and the other variables of the USA–5) are highly salient, and my interpretation of the results presented here is that hip hop artists are CONSCIOUSLY adopting the patterns of their speech community in their performances. This can be explored in future research by comparing NZ and US hip hop artists' realisation of variables that distinguish NZE and AmE but are less salient. When performers adopt their 'own accent' rather than the PSE variants, they may be doing so as an initiative act of identity (Bell 1984; Le Page & Tabouret-Keller 1985), actively trying to reduce the distance between their on-stage and off-stage personae (Coupland 2011:594).

The importance of register: Sonority and singiness

The fact that US singers themselves have such low rates of rhoticity in song no doubt relates in part to the early importance of AAE and Southern dialects in the formation of singing accent norms, as was observed by Sackett (1979). Another potential reason, however, is the preference for sonority in singing (Andres Morrissey 2008), which needs to be considered as an important non-social dimension likely to structure variation in song. Sonority has broad-reaching effects on singing styles. The AmE variants of many vowels including LOT, DRESS, and TRAP, for example, are opener, and thus more sonorous than their NZE variants, and thus have a 'sonority advantage'—the AmE variant may be preferred for both social and sonority reasons. For both BATH and nonprevocalic /r/, however, NZE has the sonority advantage over AmE: PALM is opener than TRAP, and the presence of nonprevocalic /r/ is a constriction that reduces sonority. Therefore, to use the PSE variant of LOT or DRESS is attributable to both dialect and a bias for sonority. To use the PSE variant of BATH and to produce nonprevocalic /r/, however, both involve the application of an 'Americanism', as well as going against the sonority bias, which may reinforce their heightened salience.

   Another related consideration when analysing the sociophonetics of popular music is the degree of 'singiness' in a given performance (Coddington 2004). There are likely to be systematic phonetic correlates on a cline from the most

'speaky' to the most 'singy' styles. Rap would be closer to the speaky end of this cline, with an operatic aria, for example, falling at the singiest extreme. The smaller difference between speech and the performed register in hip hop is another potential explanation for the less dramatic style-shift away from a rapper's own spoken accent in their performance. For pop singers, the register shift may be so significant that it allows the maintenance of a distinct dialect in the context of song.

Limitations

One of the most obvious weaknesses of this study is that no direct measures of spoken registers have been analysed. An ideal design would include both the speech and the singing of artists from a range of backgrounds. Another limitation is the lack of female hip hop artists in the corpus. All of the models include both male and female performers of pop music, and only male performers of hip hop. The results are all attributed to genre here but could also relate to the difference in gender between the two genres. Neither gender nor ethnicity were discussed in this article due to space limitations but are explored in greater detail in Gibson (2020).

Another limitation stems from the decision to use the Apple Music definition of artist genre in the construction of the corpus, which had both pros and cons. The reason for this choice was precisely the problem that genre divisions are notoriously fluid: pop is infused with many hip hop influences, while hip hop has moved further and further into the mainstream, and thus closer to pop. Most sources of genre information online provide multiple genre tags for any given artist, making a clear distinction between artists impractical for the kind of analysis presented in this article. The clearcut distinction provided by Apple Music shifted any bias away from me as the researcher but did introduce some classification oddities. Future work could use the tools of the rapidly evolving field of Music Information Retrieval to provide quantitative predictors sourced straight from the audio of a track. Multimodal deep learning approaches can work with features extracted from the audio along with other sources of information to classify tracks into genres (Oramas, Barbieri, Nieto Caballero, & Serra 2018). These developments present clear opportunities for the sociolinguistic study of popular music.

CONCLUSION

This study has found examples of vocal artists from both NZ and the US singing pop songs in Pop Song English, a supralocal variety which, like a standard language, appears to reduce regional and social variation. PSE is used in the restricted domain of the pop song, with several generations of music consumers having now grown up with plenty of exposure to this dialect of English. PSE has clearly defined contexts of use and well-established norms. In the context of a pop song,

it doesn't matter where the singer is from. If a pop singer WANTS their place of origin to matter, they may have to put some thought into how to sing in their 'own accent'.

For hip hop artists, there are competing discourses around projecting a 'real' self, as well as displaying membership within the Hip Hop Nation. These competing motivations lead to a diversity of rap accents, and diversity begets diversity. The marginalisation of regional variation in the phonetic styles of pop songs, however, reinforces the stability of the Pop Song English norm. In all language practice, there is a tension between convention and innovation, between centripetal and centrifugal forces (Bakhtin 1981). This tension may be particularly apparent in popular music, where performers have conflicting identities (Trudgill 1983), as a member of their speech community and as a member of the subculture associated with their musical genre. Being performed in a consistent form by singers from a range of locations, Pop Song English does not 'sound American'. The indexicalities of geographic place which would arise when hearing the same phonetic variants in a spoken interaction are backgrounded in the context of popular music. If a performer wishes to re-connect place meanings to their singing or rap, to perform 'as themselves', they must innovate away from PSE.

Whether the uniformity of PSE perseveres over the coming decades or splits into a proliferation of variation is a question worthy of sociolinguists' attention, and given the ubiquity of sung data available for analysis, the field is well placed to track its progress either way. The processes of language variation and change in popular music are likely to rely on many of the same cognitive and social processes that underlie language variation and change in traditional speech communities. However, there may be differences too, and the study of this distinct universe of variation may provide unique insights to our understanding of language in society.

APPENDIX A: SUMMARY OF BATH MODEL. LOG-ODDS OF
REALISING BATH AS TRAP (/æ /).

BATH model: TRAP ~ Genre * Country + (1 | Performer)

	ESTIMATE	STD. ERROR	Z VALUE	P VALUE
(Intercept: NZ hip hop)	−7.0218	2.1898	−3.207	0.0013
Genre: Pop	15.7971	3.431	4.604	<0.0001
Country: US	18.7575	5.2973	3.541	0.0004
Genre: Pop * Country: US	−15.9453	6.4326	−2.479	0.0132

Number of observations: 301, groups: performer, 107

APPENDIX B: SUMMARY OF RHOTICITY MODEL 1 (ALL DATA). LOG-ODDS OF REALISING NONPREVOCALIC-R.

*Rhoticity model 1 (all data): r ~ Genre * Country + NURSE + (1 + NURSE | Performer) + (1 | Word)*

	ESTIMATE	STD. ERROR	Z VALUE	P VALUE
(Intercept: NZ hip hop,not NURSE)	−2.3891	0.3158	−7.566	< 0.0001
Genre: Pop	1.0814	0.3787	2.856	0.0043
Country: US	1.6327	0.4157	3.928	< 0.0001
Lexical set: NURSE	3.2212	0.2728	11.807	< 0.0001
Genre: Pop * Country: US	−1.0108	0.512	−1.974	0.0484

Number of observations: 3,210, groups: word, 277; performer, 152

APPENDIX C: SUMMARY OF RHOTICITY MODEL 2 (US DATA ONLY). LOG-ODDS OF REALISING NONPREVOCALIC-R.

*Rhoticity model 2 (US data only): r ~ Genre * Region + NURSE + (1 | Performer) + (1 | Word)*

	ESTIMATE	STD. ERROR	Z VALUE	P VALUE
(Intercept: West/Midwest,hip hop, not NURSE)	0.1872	0.4208	0.445	0.6564
Genre: Pop	−1.1352	0.5302	−2.141	0.0323
Region: South and East	−2.1085	0.6478	−3.255	0.0011
Lexical set: NURSE	3.0976	0.3258	9.507	< 0.0001
Genre: Pop * Region:South and East	2.3288	0.7792	2.989	0.0028

Number of observations: 1,360; groups: word, 211; performer, 65

NOTES

*This research was conducted as part of a PhD project at the University of Canterbury supported by a Canterbury Scholarship. I extend my sincere gratitude to Jen Hay, Lynn Clark, and Catherine Theys for their patient and insightful supervision. I would also like to thank Robert Fromont for extensive help with corpus development and maintenance.

[1]In Gibson (2020) I used the term *standard popular music singing style* (SPMSS) for the same concept. The term *Pop Song English* clarifies that the variety under discussion refers only to English, and invites the sociolinguistic study of popular music in other languages.

[2]Vowel phonemes throughout this article are described using Wells' (1982) terms for lexical sets.

[3]Note that GOOSE was analysed as a diphthong in Gibson (2010) due to its strong dynamism. In Figure 1, the beginning of the arrow for each diphthong represents the vowel's nucleus, while the arrowhead represents the offglide. Spoken GOOSE begins at a central position and then fronts, while sung GOOSE begins at a fronted position and then retracts.

[4]For the purposes of this study, BATH also includes words in the DANCE lexical set.

[5]Variability in the BATH lexical set is a matter of phonemic alignment. In NZE, words in the BATH lexical set are realised with the same phoneme as words in the PALM lexical set (/ɑː/). Throughout

this article, I refer to NZE-type realisations as PALM, to emphasise the phonological alignment of the BATH and PALM lexical sets. In AmE/PSE, the same words are realised with the phoneme used for words in the TRAP lexical set (/æ/). I refer to these American-like realisations as TRAP. In this analysis, then, PALM and TRAP are treated as the two primary variants of the variable BATH.

[6]See http://nztop40.co.nz/chart/nzsingles.

[7]This meant some of the categorisations were of dubious accuracy. My thanks to an anonymous reviewer for pointing out some of the most obviously problematic of these (André 3000, Pharrell, and Will.I.Am were all treated as pop, while Post Malone was classified as hip hop). All of the models presented in the results section were re-run with these four artists excluded, and all of the reported results remained significant.

REFERENCES

Agha, Asif (2005). Voice, footing, enregisterment. *Journal of Linguistic Anthropology* 15(1):38–59.

Alim, H. Samy; Awad Ibrahim; & Alastair Pennycook (2009). *Global linguistic flows: Hip hop cultures, youth identities, and the politics of language*. New York: Routledge.

Andres Morrissey, Franz (2008). Liverpool to Louisiana in one lyrical line: Style choice in British rock, pop and folk singing. In Miriam A. Locher (ed.), *Standards and norms in the English language*, 195–218. Berlin: Mouton de Gruyter.

Bakhtin, Mikhail (1981). *The dialogic imagination*. Ed. by Michael Holquist. Austin: University of Texas Press.

Bates, Douglas; Martin Mächler; Ben Bolker; & Steve Walker (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1–48.

Beal, Joan 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(3):223–40.

Becker, Kara (2014). (r) we there yet? The change to rhoticity in New York City English. *Language Variation and Change* 26(2):141–68.

Bekker, Ian, & Erez Levon (2020). Parodies of whiteness: Die Antwoord and the politics of race, gender, and class in South Africa. *Language in Society* 49(1):115–47.

Bell, Allan (1984). Style as audience design. *Language in Society* 13(2):145–204.

——— (2001). Back in style: Reworking audience design. In Penelope Eckert & John R. Rickford (eds.), *Style and sociolinguistic variation*, 139–69. Cambridge: Cambridge University Press.

——— (2011). Falling in love again and again: Marlene Dietrich and the iconization of non-native English. *Journal of Sociolinguistics* 15(5):627–56. doi: 10.1111/j.1467-9841.2011.00516.x.

———, & Andy Gibson (2011). Staging language: An introduction to the sociolinguistics of performance. *Journal of Sociolinguistics* 15(5):555–72.

Best, Catherine T. (1994). Learning to perceive the sound pattern of English. In Carolyn Rovee- Collier & Lewis Lipsitt (eds.), *Advances in infancy research*, vol. 8, 217–304. Hillsdale, NJ: Ablex.

Boersma, Paul, & David Weenink (2019). *Praat: Doing phonetics by computer*. Version 6.1.04.

Brooks, William (1982). Theory and method: On being tasteless. *Popular Music* 2:9–18.

Carmichael, Katie (2017). Displacement and local linguistic practices: R-lessness in post-Katrina Greater New Orleans. *Journal of Sociolinguistics* 21(5):696–719.

Coddington, Anna (2004). *Singing as we speak? An exploratory investigation of singing pronunciation in New Zealand popular music*. Auckland: University of Auckland MA thesis.

Coupland, Nikolas (2011). Voice, place and genre in popular song performance. *Journal of Sociolinguistics* 15(5):573–602.

Cutler, Cecelia (2014). *White hip hoppers, language and identity in post-modern America*. New York: Routledge.

Duncan, Daniel (2017). Australian singer, American features: Performing authenticity in country music. *Language & Communication* 52:31–44.

Eberhardt, Maeve, & Kara Freeman (2015). 'First things first, I'm the realest': Linguistic appropriation, white privilege, and the hip-hop persona of Iggy Azalea. *Journal of Sociolinguistics* 19(3):303–27.

Flanagan, Paul J. (2019). 'A certain romance': Style shifting in the language of Alex Turner in Arctic Monkeys songs 2006–2018. *Language and Literature* 28(1):82–98. doi: 10.1177/0963947019827075.

Fromont, Robert, & Jennifer Hay (2012). LaBB-CAT: An annotation store. In Paul Cook & Scott Nowson (eds.), *Proceedings of Australasian Language Technology Association Workshop 2012*, 113–17. Dunedin: Australasian Language Technology Association.

Gerfer, Anika (2018). Global reggae and the appropriation of Jamaican Creole. *World Englishes* 37:668–83. doi: 10.1111/weng.12319.

Gibson, Andy (2005). Non-prevocalic /r/ in New Zealand hip hop. *New Zealand English Journal* 19:5–12.

Gibson, Andy (2010). *Production and perception of vowels in New Zealand popular music*. Auckland: Auckland University of Technology MPhil thesis.

———— (2011). Flight of the Conchords: Recontextualizing the voices of popular culture. *Journal of Sociolinguistics* 15(5):603–26. Online: https://doi.org/10.1111/j.1467-9841.2011.00515.x.

———— (2016). Samoan English in New Zealand: Examples of consonant features from the UC Quake-Box. *New Zealand English Journal* 29&30:25–50.

———— (2020). *Sociophonetics of popular music: Insights from corpus analysis and speech perception experiments*. Christchurch: University of Canterbury PhD thesis.

————, & Allan Bell (2012). Popular music singing as referee design. In Juan Manuel Hernández-Campoy & Juan Antonio Cutillas-Espinosa (eds.), *Style-shifting in public: New perspectives on stylistic variation*, 39–64. Amsterdam: John Benjamins.

Gilbers, Steven; Nienke Hoeksema; Kees de Bot; & Wander Lowie (2019). Regional variation in west and east coast African-American English prosody and rap flows. *Language and Speech* 63(4):713–45. doi: 10.1177/0023830919881479.

Hermastuti, Sendy Intania, & Arina Isti'anah (2018). Consonant changes in Korean singers' pronunciation. *Journal of Language and Literature* 18(1):28–35. doi: 10.24071/joll.v18i1.1050.

Hess, Mickey (2009). *Hip hop in America: A regional guide*. Santa Barbara, CA: Greenwood Press.

Jansen, Lisa (2018). 'Britpop is a thing, damn it': On British attitudes toward American English and an Americanized singing style. In Valentin Werner (ed.), *The language of pop culture*, 116–35. New York: Routledge.

————, & Michael Westphal (2017). Rihanna works her multivocal pop persona: A morpho-syntactic and accent analysis of Rihanna's singing style. *English Today* 33(2). doi: 10.1017/s0266078416000651.

Konert-Panek, Monika (2017a). Americanisation versus Cockney: Stylisation in Amy Winehouse's singing accent. In Victor Kennedy & Michelle Gadpaille (eds.), *Ethnic and cultural identity in music and song lyrics*, 77–94. Newcastle Upon Tyne: Cambridge Scholars.

———— (2017b). Overshooting Americanisation: Accent stylisation in pop singing – Acoustic properties of the bath and trap vowels in focus. *Research in Language* 15(4):371–84.

———— (2018). Singing accent Americanisation in the light of frequency effects: LOT unrounding and PRICE monophthongisation in focus. *Research in Language* 16(2):155–68. doi: 10.2478/rela-2018-0008.

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

Lenth, Russell (2020). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.5.2-1.

Liberman, Alvin M.; Katherine Safford Harris; Howard S. Hoffman; & Belver C. Griffith (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5):358–68.

Marsden, Sharon (2017). Are New Zealanders 'rhotic'? *English World-Wide* 38(3):275–304. doi: 10.1075/eww.38.3.02mar.

Mitchell, Tony (2008). Doin' damage in my native language: The use of 'resistance vernaculars' in hip hop in France, Italy, and Aotearoa/New Zealand. *Popular Music and Society* 24(3):41–54. doi: 10.1080/03007760008591775.

Murphey, Tim (1992). The discourse of pop songs. *TESOL Quarterly* 26(4):770–74. Online: https://doi.org/10.2307/3586887.

Nosek, Brian; Charles Ebersole; Alexander DeHaven; & David Mellor (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences* 115(11):2600–2606.

O'Hanlon, Renae (2006). Australian hip hop: A sociolinguistic investigation. *Australian Journal of Linguistics* 26(2):193–209.

Oramas, Sergio; Francesco Barbieri; Oriol Nieto ssCaballero; & Xavier Serra (2018). Multimodal deep learning for music genre classification. *Transactions of the International Society for Music Information Retrieval* 1(1):4–21.

Pennycook, Alastair, & Tony Mitchell (2009). Hip hop as dusty foot philosophy: Engaging locality. In H. Samy Alim, Awad Ibrahim, & Alastair Pennycook (eds.), *Global linguistic flows: Hip hop cultures, youth identities, and the politics of language*, 25–42. New York: Routledge.

Pichler, Pia, & Nathanael Williams (2016). Hipsters in the hood: Authenticating indexicalities in young men's hip-hop talk. *Language in Society* 45(4):557–81.

R Core Team (2019). R: A language and environment for statistical computing, version 3.6.1. Online: https://www.R-project.org.

Sackett, Samuel J. (1979). Prestige dialect and the pop singer. *American Speech* 54(3):234–37. doi: 10.2307/454954.

Simpson, Paul (1999). Language, culture and identity: With (another) look at accents in pop and rock singing. *Multilingua* 18(4):343–67.

Thomas, Erik (2003). Rural white southern accents. In Bernd Kortmann & Edgar W. Schneider (eds.), *A handbook of varieties of English: A multimedia reference tool*, 300–24. New York: Mouton de Gruyter.

Trudgill, Peter (1983). *On dialect: Social and geographical perspectives*. Oxford: Blackwell.

Villarreal, Dan; Lynn Clark; Jennifer Hay; & Kevin Watson (2021). Gender separation and the speech community: Rhoticity in early 20th century Southland New Zealand English. *Language Variation and Change* 33(2):245–66. doi: 10.1017/S0954394521000090.

Watts, Richard J., & Franz Andres Morrissey (2019). *Language, the singer and the song: The sociolinguistics of folk performance*. Cambridge: Cambridge University Press.

Wells, John C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Werker, Janet F., & Takao K. Hensch (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology* 66:173–96.

Westphal, Michael (2018). Pop culture and the global spread of non-standardized varieties of English: Jamaican Creole in German reggae subculture. In Valentin Werner (ed.), *The language of pop culture*, 95–115. New York: Routledge.

———, & Lisa Jansen (2021). English in global pop music. In Britta Schneider, Theresa Heyd, & Mario Saraceni (eds.), *Bloomsbury World Englishes, vol. 1: Paradigms*, 190–206. London: Bloomsbury Academic.

Wickham, Hadley (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer.

Williams, Quentin (2017). *Remix multilingualism: Hip hop, ethnography and performing marginalized voices*. London: Bloomsbury.

Wilson, Guyanne (2017). Conflicting language ideologies in choral singing in Trinidad. *Language & Communication* 52:19–30. doi: https://doi.org/10.1016/j.langcom.2016.08.003.

Wolfram, Walt, & Erik Thomas (2002). *The development of African American English*. Oxford: Blackwell.

Wray, Alison (1999). Singers on the trail of 'authentic' Early Modern English: The puzzling case of /æ:/ and /ɛ:/. *Transactions of the Philological Society* 97(2):185–211.

Yaeger-Dror, Malcah (1991). Linguistic evidence for social psychological attitudes: Hyperaccommodation or (r) by singers from a Mizrahi background. *Language and Communication* 11(4):309–31.

——— (1993). Linguistic analysis of dialect 'correction' and its interaction with cognitive salience. *Language Variation and Change* 5(2):189–224.

Zhou, Sijing, & Andrew Moody (2017). English in the voice of China. *World Englishes* 36(4):554–70. doi: 10.1111/weng.12240.

A N D Y   G I B S O N

Address for correspondence:
Andy Gibson
Centre for Language Sciences
Macquarie University, Sydney
16 University Avenue, Macquarie University
NSW 2109, Australia
gibsonism@gmail.com