Routledge
Taylor & Francis Group

Australian Journal of Linguistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cajl20

Ethnolectal and community change ov(er) time: Word-final (er) in Australian English

James Grama , Catherine E. Travis & Simon Gonzalez

To cite this article: James Grama , Catherine E. Travis & Simon Gonzalez (2020) Ethnolectal and community change ov(er) time: Word-final (er) in Australian English, Australian Journal of Linguistics, 40:3, 346-368, DOI: 10.1080/07268602.2020.1823818

To link to this article: https://doi.org/10.1080/07268602.2020.1823818

AUSTRALIAN JOURNAL OF LINGUISTICS
2020, VOL. 40, NO. 3, 346–368
https://doi.org/10.1080/07268602.2020.1823818

Ethnolectal and community change ov(er) time: Word-final (er) in Australian English

James Grama, Catherine E. Travis and Simon Gonzalez

University of Duisburg-Essen, Australian National University and Australian National University

ABSTRACT

Increased global migration to international urban centres has motivated a growing interest in ethnolects and the role migrant communities play in language variation and change. Here, we consider ethnolectal variation in real and apparent time, by examining the realization of word-final (er) (e.g. *teacher*, *remember*) in Australian English. We capitalize on sociolinguistic interview data collected by Barbara Horvath in Sydney in the 1970s as a benchmark against which to compare newly collected recordings with Sydneysiders in the 2010s. Approximately 15,000 tokens of word-final (er) were extracted from the speech of nearly 200 people, including Anglo-Australians, and second-generation migrants of Italian, Greek and Chinese background. Acoustic analyses of vowel duration and position in the vowel space reveal incremental lengthening with concomitant lowering and backing over time for (er), though only in prosodically final position. This change was led by Greek and Italian teenagers in the 1970s, then taken up by working class women, and today, has been adopted across the community. Tracking this change in real and apparent time provides evidence that ethnolectal features may be adopted by the wider community, with ethnic minorities playing a leading role in language change.

ARTICLE HISTORY
Accepted 1 September 2020

KEYWORDS
Ethnolects; sociophonetics; word-final (er); Australian English; language change; social conditioning

1. Ethnolectal variation

The role played by ethnicity in language variation has been recognized since early variationist research (e.g. Labov, 1966, 1972). More recently, research on ethnolectal variation has come to the fore, sparked by increased mobility worldwide which has given rise to extensive international migration and greater diversity in major urban centres, for example, in England (e.g. Cheshire et al., 2011), Canada (e.g. Hoffman & Walker, 2010) and Sweden (e.g. Gross et al., 2016). *Ethnolects* are understood to be new ways of speaking, associated with second- or third-generation migrants who are native speakers of the majority language, and may not speak the community language. Ethnolectal features cannot be directly attributed to second language transfer (Labov, 2008), and are considered to serve "as a means of establishing a distinctive linguistic identity" (Clyne et al., 2002, p. 134).

CONTACT James Grama james.grama@uni-due.de

In Australia, pioneering work by Michael Clyne and colleagues presented overviews of ethnolectal Greek, Yiddish and German Australian English, seen in a combination of lexical, phonological and morphosyntactic features (Clyne et al., 2001, 2002). Quantitative analysis has reported ethnolectal use of specific linguistic features, such as the fronting of /θ/ and deletion of word-initial /h/ for Italian-Australians (Horvath, 1985, pp. 102–103), and realization of /l/, voice onset time and prosody for Lebanese-Australians (Clothier, 2019; Clothier & Loakes, 2018; Cox & Palethorpe, 2011) (see Clothier, 2020 for a review). Other studies, however, have found little evidence of ethnolectal variation. In sociolinguistic interview data with Anglo-, Greek- and Italian-Australian teenagers recorded in 1970s Australia, Horvath found similar realizations of diphthongs for these three groups (though with differences in social conditioning) (1985, pp. 80–82), and Guy and colleagues found similar patterning for high rising terminal (Guy et al., 1986, p. 40).

Though Horvath found the diphthong realizations to be qualitatively similar across ethnic groups, the distribution of the realizations indicated that teenage Greek- and Italian-Australians were leading a change towards more 'general' Australian English vowel realizations, away from both 'broad' and 'cultivated' speech, respectively associated with the working and middle class (1985, pp. 91–94). She interpreted this as young ethnic minorities trying to 'sound Australian' by differentiating themselves from their first-generation migrant parents, whose accented vowels were situated at the broad or cultivated ends of the continuum (Horvath, 1985, pp. 75–76). Other apparent-time studies have lent support to the notion that ethnic minorities may lead in change (e.g. Cheshire et al., 2011; Gross et al., 2016).

In this paper, we present one of the first real-time studies of ethnolectal variation, through an examination of word-final (er) – unstressed /ə/ in word-final position in words with a following *r* in the orthography, such as *teacher*, *remember* and *culture*. A distinct realization of (er) has been associated with ethnolectal variation. Warren refers to "widespread use of [a] in final syllables as in 'my pleasure'" (1999, p. 92), and Clyne et al. describe this as one of "the most conspicuous phonological features of the [Greek] ethnolect" (2001, p. 228). Subsequent acoustic analyses indicate a longer and lower and/or backer realization of (er) by Greek-, Lebanese- and Italian-Australians compared with Anglo-Australians in the 1990s (Kiesling, 2005), and by Lebanese-Australians in the 2000s (Clothier, 2014). Based on apparent-time comparisons of younger and older speakers, Kiesling hypothesized that Greek-Australians in the 1990s may have been leading a change towards a longer and backer (er) across the community (2005, p. 33).

To test this in real time, we compare the realization of (er) in 1970s Australia with that in the 2010s across different ethnic groups, namely Anglo-Australians, and second-generation migrants of Greek, Italian and Chinese background (referred to here as Greek-, Italian- and Chinese-Australians). We begin in §2 with an overview of ethnic diversity in Australia, and how this has changed over the past four decades. In §3 we outline the data, comprising sociolinguistic interviews from nearly 200 Australians, from recordings made in the 1970s and the 2010s. From these data, nearly 15,000 (er) tokens were extracted for analysis, as described in §4. Section 5 reports the changes observed, which involve a lengthening of (er) over time, first in the speech of Greek- and Italian-Australians, then in that of working class females, before being more widely adopted by Anglo- and Italian-Australians, and, to a lesser degree, Chinese-Australians. We conclude in §6, highlighting that ethnolectal features need not be stable over time, but can be

adopted by the wider community, and we consider some possible accounts of why ethnic minorities might be at the forefront of language change.

2. Ethnic diversity in Australia

Despite being a predominantly English-speaking country, Australia is highly ethnically and linguistically diverse. According to the most recent national census in 2016, over one-fifth of Australians use a language other than English in the home; over 300 distinct languages are spoken (including migrant and indigenous languages); and nearly one half of the population was either born overseas, or has at least one parent born overseas (Australian Bureau of Statistics, 2016b).

The makeup of migrants to Australia has become more diverse over time. In the post-World War II era, Australia saw large-scale migration from eastern and central Europe, and in the 1970s, the lifting of the White Australia Policy (which had restricted immigration to people of European origin) opened Australia to immigration from Asia, the Middle East and Africa (cf. Clyne & Kipp, 1997; Jupp, 2001; Willoughby & Manns, 2020). This demographic change has had tangible linguistic impacts – in 1981, just 12% of the population spoke a community language at home, compared with 21% today (Australian Bureau of Statistics, 1981, 2016b). The nature of that diversity has also changed, seen, for example, in the most widely spoken languages. Figure 1 shows the community languages most widely spoken in Australia in 1991 (the earliest year for which comparable census data are available), and in 2016 (the most recent census at time of writing), with the proportion of the population of community-language speakers represented in each year.* The most widely spoken community languages in 1991 were Italian and Greek (from the post-World War II migration schemes), followed distantly by Cantonese (primarily from a Business

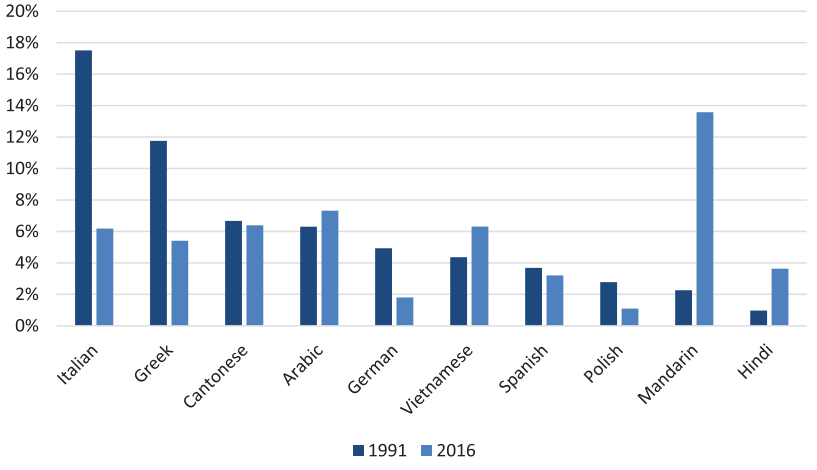

Figure 1 Most widely spoken community languages in Australia in 1991 and in 2016 as a percentage of the total number of community-language speakers (Australian Bureau of Statistics, 1991, 2016b)

* Seven of the ten languages shown here were among the eight most widely spoken in both years; of those that were not, German and Polish were in the top eight in 1991 only, and Hindi was in the top eight in 2016 only.

Migration programme conducted in the 1980s with Hong Kong (cf. Jupp, 2001, p. 218)). Today, the proportion of Italian and Greek speakers has dropped, while that of Cantonese speakers has remained stable. Each of these three languages now accounts for a far lower proportion of community-language speakers than Mandarin does, which is overwhelmingly the most widely spoken community language.

For this study, we focus on second-generation migrants of Greek, Italian and Chinese (specifically, Cantonese-speaking) background, who grew up in Australia as speakers of Australian English. These are often the children of those represented in Figure 1, as second-generation migrants do not necessarily speak the community language. Thus, Figure 1 only indirectly captures the size of the communities under study here, but it does give an indication of the history of these migrant groups in Australia – Greek and Italian as well-established migrant groups, and Chinese as a newer, and younger community – which is also reflected in the participant profile of the data.[1]

3. Data for examining ethnolectal change over time: The *Sydney Speaks* corpus

Sydney Speaks is a large-scale sociolinguistic project examining variation and change over time and across diverse social groups (Travis, 2016–2021). The project is centred in Sydney which, as Australia's largest and most diverse city, is an ideal location to study the linguistic impact of increasing diversity. Out of its some five million inhabitants, close to two-fifths (38%) report speaking a community language at home, and two-thirds were either born overseas, or have one or both of their parents born overseas (Australian Bureau of Statistics, 2016b). Given the relative homogeneity of English across major urban centres in Australia (cf. Cox & Palethorpe, 2012, p. 299), we predict that the results here would generally hold in other large Australian cities, and this is supported by observations about realizations of (er) in English spoken in Melbourne (Clothier, 2014; Clyne et al., 2001, p. 228; Warren, 1999, p. 92).

The project brings together sociolinguistic interview data recorded at different time periods and with different age groups, allowing for the tracking of change in real and apparent time. Participants are all native speakers of Australian English and are stratified according to age, sex, socioeconomic status and ethnicity. Particularly novel in this project is the capacity for longitudinal study of ethnic variation and its role in language variation and change in the wider community.

3.1 Participants

We utilize two corpora for the current analysis: the *Sydney Social Dialect Survey*, recorded in 1977–1981 (Horvath, 1985), and the *Sydney Speaks* 2010s corpus, recorded some 40 years on, from 2016 to the present, and for which data collection is ongoing (Travis et al., in progress). All participants were living in Sydney at the time of recording, and

[1] The ethnic makeup of second-generation migrants cannot be readily ascertained from census data. The Australian census asks about ethnicity for Aboriginal and Torres Strait Islander people, but does not ask specifically about ethnicity more broadly. It also does not ask about parents' language, nor (from 2001 onwards) parents' country of birth. While it does ask about ancestry, the responses to this question are difficult to interpret, and the ABS recommends considering these results alongside language, religion and country-of-birth data (Australian Bureau of Statistics, 2016a).

Table 1 Participants by age, sex, ethnicity and socioeconomic status

	1970s				2010s				
	Adults b. 1930s 32–64 y/o		Teens b. 1960s 12–18 y/o		Adults b. 1960s 42–61 y/o		Young Adults b. 1990s 18–31 y/o		Total speakers/ ethnicity & class
	Female	Male	Female	Male	Female	Male	Female	Male	
Anglos									
Middle	4	2	5	4	4	4	4	4	31
Upper Working	4	3	3	4	4	4	5	4	31
Lower Working	4	3	4	4	3	2	3	2	25
Greeks									
Middle			2	3					5
Upper Working			3	5					8
Lower Working			5	5					10
Italians									
Middle			5	2	4	5	3	1	20
Upper Working			3	4	4	3	2	2	18
Lower Working			5	6	2	4	3	3	23
Chinese									
Middle							6	7	13
Upper Working							5	4	9
Total speakers / age & sex	12	8	35	37	21	22	31	27	193

the 2010s participants were also all born in Sydney, or moved there before the age of six, and had spent all or most of their lives there. A total of 193 participants are included, stratified according to age, sex, ethnicity and socioeconomic status, as summarized in Table 1.

Participants represent four age groups across real and apparent time: Adults and Teenagers from the 1970s, and Adults and Young Adults from the 2010s. These four age groups span 60 years in apparent time – the 1970s Adults were born around the 1930s; the 1970s Teens and 2010s Adults around the 1960s (corresponding to the same generation); and the 2010s Young Adults around the 1990s.

Four ethnic groups are included: Anglo, Greek, Italian and Chinese. We use the term 'Anglo' to refer to Australians of Anglo-Celtic background. All were born and raised in Australia and, for the 2010s, are minimally fourth-generation Australians, that is, their parents and at least three grandparents were born and raised in Australia.[2]

For the ethnic minorities, we focus on second-generation migrants, which we define as those who were either born in Australia or arrived before the age of six (prior to beginning school), and whose parents migrated to Australia as adults. The groups included were determined both by migration history and progress in data collection. There was a sizable community of young second-generation Greek- and Italian-Australians in the 1970s, and thus teenagers of these two groups are included. We are not yet able to report on patterns for 2010s Greek-Australians, as data collection is still underway. The Italian-Australian community today is an older community, and thus we include second-generation 2010s Adults, and for Young Adults, second- and third-generation migrants.[3] The Chinese-background community, on the other hand, is a newer community, and thus is represented in the sample only for the Young Adult age group. The

[2] Six of the 43 Anglo participants in the 2010s corpus included here report having one grandparent who was born outside Australia in another English-speaking country; one participant reports having an Aboriginal-Australian grandparent. We do not have information about the grandparents for the 1970s participants.

[3] Of the 14 Young Adult Italians, nine are second-generation and five are third-generation migrants.

parents of the participants are Cantonese-speaking migrants primarily from Hong Kong, but also from Guangzhou in mainland China. We refer to these participants as Chinese-Australians, consistent with how they describe themselves in the sociolinguistic interviews, and noting that not all speak Cantonese.

Socioeconomic class has not been widely studied in work on Australian English, and there are no agreed-upon metrics for classification. In the *Sydney Speaks* project, we face the additional complexity of establishing meaningful socioeconomic groups that are comparable across different time periods and ethnic groups. In the analyses presented here we categorize participants using four social variables that have been widely used for class assignations in sociolinguistics: occupation, suburb, education level and school type. To establish these categories, each participant was initially given a score for each variable. For occupation (the participant's, or that of their father, mother or spouse where relevant, e.g. for 1970s Teens), we used the Australian Socioeconomic Index, AUSEI06 (McMillan et al., 2009); for suburb, we used the Index of Relative Socio-economic Disadvantage assigned by the Australian Bureau of Statistics to each suburb for 2016, or postcode for 1986 (Australian Bureau of Statistics, 1986, 2016b); and for education level, we used an adaptation of the Australian Qualifications Framework (2013). For school type, three levels were scored as a measure of relative privilege based on the funding model – government, catholic or private. For 2010s Young Adults only, we used the Index of Community Socio-Educational Advantage scores from 2016 provided on the *My School* website (Australian Curriculum Assessment and Reporting Authority, 2019). These measures apply differently according to time period, ethnicity and, in some cases, sex. For example, the 1970s Adults did not have the same access to education that 2010s Adults had, and the 2010s Chinese-Australian Young Adults do not cover the same range of occupations as their Anglo and Italian peers. We therefore sought to establish rankings that were meaningful for each community. To do this, we conducted Hierarchical k-means cluster analyses (Hartigan & Wong, 1979), using *factoextra* (Kassambara & Mundt, 2020) in R (R Core Team, 2018), for each age group and ethnicity, and, for the 1970s Adults, for each sex. Within each analysis, participants were hierarchically nested according to their scores across the four variables, cluster centres were calculated and, based on their Euclidean Distance from cluster centres and from other members, participants were assigned to a pre-specified number of clusters which we interpret as corresponding to socioeconomic status. For all but the Young Chinese-Australians, we specified three clusters. Following Horvath (1985, p. 47), we term these Lower Working, Upper Working and Middle Class, though these labels are best understood relative to each other, rather than as values with an absolute meaning. Because the sample of Young Chinese-Australians does not cover a comparably wide social spread, we specified two clusters, which correspond most closely to Upper Working and Middle Class in the other groups.

It is worth providing some more information about the social makeup of these communities in the Australian context. Greek- and Italian-Australians share a similar migration history and, from participants' self-reports, similar networks and lived experiences in the context of a chiefly white Australia. Both groups form part of the social construct represented by the term "'wog' […] an ethnic slur in Australian English denoting Australians of Mediterranean or Southern European background" (Sala et al., 2010, p. 114). Many participants recount experiencing discriminatory uses of this term, but they also talk of its positive use in in-group interactions today, evidencing a degree of reappropriation (as

noted, for example, by Kiesling, 2005, p. 5; Sala et al., 2010, p. 114; Warren, 1999, p. 90). This in-group use is reflected in what Warren terms "wogspeak", which she reports is positively evaluated by actors showcasing this variety in the media in the 1990s (1999, pp. 90–91). "Wogspeak" is described, by one of the actors, as a "real working class ethnic accent" (1999, p. 91), highlighting class associations to which we will return.

The Chinese-Australians studied here, on the other hand, tend to orient to higher socioeconomic status, reflected both in their linguistic patterning (e.g. for vowels (Grama et al., to appear)), and in their social profile. Recent census reports indicate that second-generation Chinese-Australians tend to be employed in high-status occupations, be well educated, attend selective or private schools, and live in wealthier suburbs (Australian Bureau of Statistics, 2016b). This community also has an ethnic label, again, mentioned by many of our participants, that of 'ABC', or Australian-Born Chinese. Unlike that used for the Greeks and Italians, however, this is an in-group term that is not in wide use outside the community.

3.2 Spontaneous speech data

Spontaneous speech was recorded via sociolinguistic interviews, which take the form of an informal interview aimed at getting the participant talking as naturally as possible, paying minimal attention to their speech (Labov, 1984, pp. 32–42). While the 1970s interviewers were typically adults with some connection with the lead researcher (Barbara Horvath) and the university environment, the 2010s interviews were conducted by community members, who generally recorded people they knew, including friends, family members or others drawn from their extended network. Familiarity between the interviewer and the participant has been noted to enhance the naturalness of the speech, and thus the possibility of recording ethnolectal features, which may be less likely to occur in conversations with outsiders (cf. Clyne et al., 2002). We might then expect greater ethnic differences among the 2010s participants, but, as we will see, this is not the case.

The 1970s interviews lasted around 45 minutes, and the 2010s interviews from 60 to 90 minutes. We selected approximately 30 minutes (or 5,000 words of speech) to transcribe from each speaker, providing close to 100 hours and one million words of speech for the analyses presented here.

3.3 Prosodic transcription

We produced time-aligned orthographic transcriptions in ELAN (Lausberg & Sloetjes, 2009), breaking the data into stretches of speech clearly delineated by pauses. These transcriptions were force aligned in LaBB-CAT (Fromont & Hay, 2012), after which acoustic measurements of word-final (er) were taken (see §4).

Following forced alignment, transcriptions were further polished in ELAN, with the addition of prosodic information in accordance with standardized protocols (outlined in Du Bois et al., 1993). Central to this method is the Intonation Unit (IU), described as "a stretch of speech uttered under a single, coherent intonation contour" (1993, p. 47). IUs generally correspond to intonational phrases (Pierrehumbert & Hirschberg, 1990), but they are defined on the basis of their boundaries, rather than their internal structure.

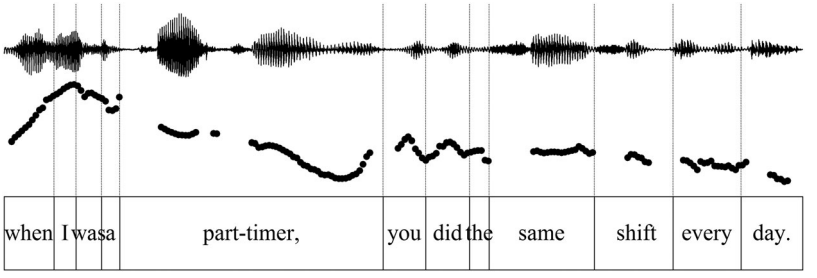

Figure 2 Pitch track (in Hz) and waveform of two Intonation Units with continuing and final intonation

(1)
Parker: *when I was a part-timer,*
you did the same shift every day.
[SydS_AOM_105: 50:43-50:47][4]

They are identified auditorily, drawing on a composite of features, including pausing between IUs; pitch reset (or a perceived rise in pitch) and anacrusis at the beginning of an IU; and lengthening at the end of an IU (Chafe, 1994, pp. 58–60). These features can be seen in Figure 2, which presents the waveform and pitch track of the two IUs in example (1). Here, we see faster speech early in the IU and lengthening towards the end, as well as higher pitch at the onset of both IUs (on 'when I' and 'you', respectively), following which the pitch gradually drops over the course of the IU. The prosodic transcription allows the nature of the change in (er) over time to be identified because, as we observe below, the realization of (er) is strongly tied to IU position.

Also central to this method is the notion of transitional continuity, "the degree of continuity that occurs at the transition point between one intonation unit and the next" (Du Bois et al., 1993, p. 53). Transitional continuity falls into one of three categories: 'continuing' intonation, which refers to a range of contours including a slightly rising, level or slightly falling pitch (indicated by a comma; illustrated in the first line in (1)); 'final' intonation, characterized by a fall to low pitch (indicated by a full stop; seen in the second line in (1)); and 'appeal' intonation, a high rise in pitch (indicated by a question mark; seen in (2) and Figure 3) (Du Bois et al., 1993, pp. 54–55). Truncated IUs, where the speaker breaks off before completing the prosodic contour, are marked with a double hyphen; due to the heterogeneity of this category, (er) tokens occurring IU-finally with truncated IUs are excluded from the analyses below ($n = 326$).

(2)
Isa: *I had an English teacher?*
[SSDS_GTF_855: 27:46–27:48]

4. Extracting (er) for analysis

We define word-final (er) as unstressed /ə/ in word-final position in minimally disyllabic words with a following *r* in the orthography. The most common orthographic

[4] In brackets following each example are the corpus name, the speaker social code (community, age group, sex) and speaker number, and time stamps of the beginning and end of the excerpt. All names given are pseudonyms.

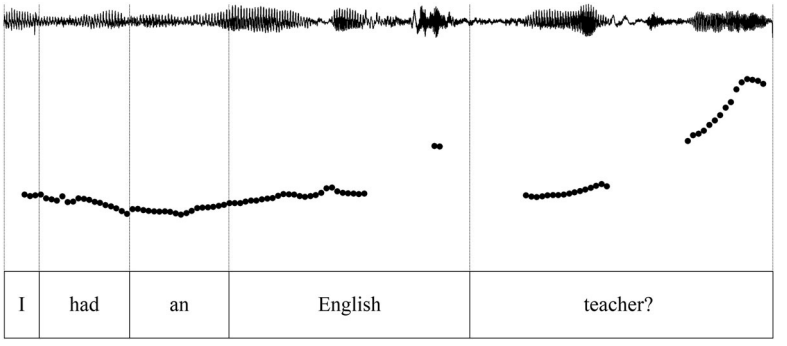

Figure 3 Pitch track (in Hz) and waveform of an Intonation Unit with appeal intonation

representation of this form is *-er*, as in *teacher*, *remember*, *whatever*. We also include other spellings (e.g. *culture*, *colour*, *particular*, *centre*), which together account for under 10% of all (er) tokens in the data.

Analyses were based on the automatic alignment produced by LaBB-CAT.[5] Vowel duration and midpoint F1 and F2 measurements were extracted using the *rPraat* (Bořil & Skarnitzl, 2016) and *PraatR* (Albin, 2014) packages in R. As measurements from only a single vowel were available, formant values were normalized using the vowel-intrinsic approach developed by Bladon et al. (1984) (as recommended by Flynn and Foulkes (2011)).

Each instance of (er) extracted from the force-aligned transcriptions was matched (via time stamps) with the corresponding instance in the prosodic transcription, to obtain information about IU-position and transitional continuity (which was not included in the transcription in LaBB-CAT). This process yielded a total of 14,872 tokens for which both acoustic and prosodic information were available (an average of 77 tokens per speaker, ranging from 12 to 165).[6]

Analyses were conducted by fitting separate linear mixed-effects models using *lme4* (Bates et al., 2015), with the duration of (er) as the dependent variable. All models tested for logical two- and three-way interactions between independent variables, and included random intercepts for speaker and word. Model fit was assessed via ANOVA; *p*-values from linear mixed-effects models were derived using *sjPlot* (Lüdecke & Schwemmer, 2020). Models that were used to evaluate changes in word-final (er) over time included age as a predictor (see §5.1). Models used to probe social conditioning reported in §5.3 and §5.4 were fit to each age cohort separately, and initially included community, sex and class in a three-way interaction; interactions that did not reach significance were simplified to include predictors as main effects, though all factors were retained to facilitate comparisons across groups.

Speech rate was included in all models to ensure the reliability of any durational differences observed. This was calculated in syllables per minute, determined on the basis of

[5] Systematic checking was deemed unnecessary due to the high quality of boundary placement by LaBB-CAT (cf. Gonzalez et al., 2020), as confirmed by spot checking of the (er) alignment.

[6] Approximately 20% of the (er) tokens in the prosodic transcription were not matched with the corresponding LaBB-CAT tier (due to differences across transcription versions in LaBB-CAT and ELAN, for example when the timestamps were not identical), and were excluded from analysis.

the number of word-level phonemic representations in the time interval in which the (er) token occurred (these time intervals being the stretches of speech occurring between pauses in the transcriptions entered into LaBB-CAT).[7] Speech rate impacted duration for all models run on the data in the predicted direction – faster speech rate correlated with significantly shorter (er). Including speech rate in the models allows us to establish that the changes in duration we observe are not merely artefacts of speech rate, but rather genuine changes, whereby (er) has lengthened over time.

5. Change ov(er) time

5.1 Identifying the nature of the change

Phrase-final lengthening is a well-established trait of final syllable rhymes in English (e.g. Klatt, 1976), and is one of the characteristics of IU-final position, as noted above. In line with this, we observe that IU position has a strong impact on the realization of (er), as (er) is significantly longer in IU-final than in medial position. This is evident in Figure 4, which shows the raw measurements in milliseconds for the duration of (er) by IU position for all speakers, at each of the four time points. We see here that, not only is IU-final (er) longer than IU-medial (er), it increases substantially more over time, with (er) going from a mean of 77 ms for 1970s Adults to 118 ms for 2010s Young Adults in IU-final position, compared with an increase from 53 ms to 63 ms in IU-medial position.

IU-final lengthening over time is corroborated by a linear mixed-effects model fit to (er) duration, with IU position, age and speech rate as predictors, and an interaction between IU position and age, given in Table 2. While speech rate has the expected inhibiting effect on the duration of (er), this model confirms that lengthening over time is not due to a change in speech rate over time. Here we see that, though there is some lengthening for IU-medial (er) across corpus (such that 2010s speakers exhibit significantly longer IU-medial (er) than 1970s speakers), the change is greater in final position. This is reflected in the larger estimates in the interaction between IU-final position and age (capturing the degree of lengthening over time in IU-final position) than by age alone (capturing lengthening in IU-medial position). Furthermore, it is in the estimates for this interaction (that is, in IU-final position only) that we observe an incremental increase across each age group. Relevelled models, with 1970s Teens and with 2010s Adults as reference levels, confirm that the change for IU-final (er) progresses incrementally across age group, with 2010s Adults significantly longer than 1970s Teens ($\beta = 9.61$, $df = 12{,}689$, $p = 0.001$) and 2010s Young Adults longer than 2010s Adults ($\beta = 6.67$, $df = 12{,}689$, $p = 0.022$). This allows us to establish that the main change has been in IU-final position. Failing to separate IU-final from IU-medial contexts may mask the progress of this change, as two-thirds of the data occur in medial position.

We also note that there appears to be a general inhibiting effect of following segment, as, even in IU-final position, (er) is shorter in words that contain a final /s/ (e.g. *teachers*, *remembers*, $n = 898$). Furthermore, in IU-final position, (er) undergoes less lengthening

[7] Speech rate was calculated based on pause-to-pause intervals in LaBB-CAT rather than IUs because of the convention in the prosodic transcription method of assigning all of a sound file (including silences) to an IU, meaning that pauses are included in IUs, rendering them less reliable as a measure of speech rate.

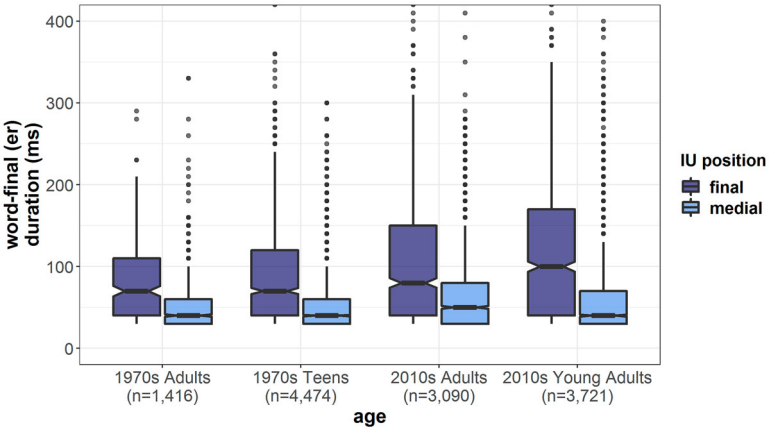

Figure 4 Duration of (er) for all speakers over time in IU-final ($n = 3{,}423$) vs. medial ($n = 9{,}278$) position

Table 2 Output of linear mixed effects fit to (er) duration for all speakers ($n = 12{,}701$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=1970s Adults, IU medial]	85.85	3.85	22.32	**<0.001**
IU final	24.01	3.56	6.74	**<0.001**
1970s Teens	-1.43	3.80	-0.38	0.707
2010s Adults	13.27	4.05	3.28	**0.001**
2010s Young Adults	12.49	3.90	3.20	**0.001**
Speech rate	-6.58	0.31	-20.90	**<0.001**
IU final: 1970s Teens	11.24	4.00	2.81	**0.005**
IU final: 2010s Adults	20.86	4.16	5.02	**<0.001**
IU final: 2010s Young Adults	27.53	4.04	6.81	**<0.001**

Note: Excluded here are IU-final tokens with truncated intonation (§3.3, $n = 326$), and tokens with a final /s/ (this section, IU-medial $n = 947$, IU-final $n = 898$).

over time with a final /s/ than without (with final /s/: 1970s Adult mean = 74 ms, 2010s Young Adult mean = 96 ms; compared with the lengthening from 77 ms to 118 ms without final /s/ reported above).[8]

A further consideration is transitional continuity, that is, whether the IU ends with continuing, final or appeal intonation. Comparison of these three reveals that (er) tends to be shorter with final intonation (mean = 87 ms, $n = 1{,}431$) than continuing (mean = 105 ms, $n = 1{,}753$) and appeal intonation (mean = 126 ms, $n = 239$). Though we have not examined high rising terminals in these data, the category would overlap with that of appeal intonation, and thus this patterning is consistent with Kiesling's observation that longer (er) is associated with high rising terminals (2005, p. 19). A

[8] The general inhibiting effect of following segment is evident in IU-medial position, where the marginal lengthening observed is primarily due to the small number of IU-medial pre-pausal tokens ($n = 86$), which go from a mean of 101 ms for 1970s Adults to 176 ms for 2010s Young Adults. Interestingly, IU-medial pre-vocalic tokens (e.g. *remember it*, *whatever else*) lengthen similarly to pre-consonantal tokens (pre-vocalic means: 51 ms 1970s Adults; 61 ms 2010s Young Adults; pre-consonantal means: 51 ms 1970s Adults; 57 ms 2010s Young Adults). This appears to be due to hiatus resolution, which, based on auditory observations on a subset of the data, includes preservation of /ɹ/, but also glottalization, and glide insertion.

linear mixed-effects model fit to IU-final (er) tokens with an interaction between transitional continuity and age (presented in Table A1 in the Appendix) reveals that while (er) is longer with both continuing and appeal intonation than with final intonation, this effect is significant for continuing intonation only (no doubt due to the small number of instances of appeal intonation, meaning that further exploration must be left for future research).[9] This model also shows that the effect for transitional continuity is consistent across age group (evidenced by the lack of significant interaction in the model). Thus, we conclude that the lengthening observed is not due to any single intonation contour, and we do not distinguish between the different transitional continuities for the purposes of this paper.

Finally, lengthening is not impacted by morphological status, that is, whether (er) constitutes a distinct morpheme (e.g. *teach-er*, *old-er*) or not (e.g. *remember*, *whatever*). This null-effect is confirmed by the model presented in Table A1 in the Appendix.

Thus, we focus here on (er) in open syllables in IU-final position (the data presented in Figure 4 and Table 2), regardless of transitional continuity or (er)'s morphological status.

5.2 Duration of (er) and vowel quality

Vowel duration impacts position in the vowel space, such that longer vowels tend to be more peripheral, and shorter vowels more central (Fletcher & McVeigh, 1993, pp. 360–361; Liljencrants & Lindblom, 1972). As a result, we might expect lengthening over time to coincide with lowering and retraction of (er) and this is indeed the case, as depicted in Figure 5.

We observe two things from this figure. First, longer tokens (marked in darker shades) become more frequent over time, such that the 2010s Young Adults exhibit the highest proportion of longer tokens. This is consistent with the lengthening captured in Figure 4. Second, there is a clear shift in the concentration of (er) tokens in the vowel space over time, such that each subsequent generation exhibits a lower and more retracted (er) cluster. As we might predict, the lowest and backest tokens tend to be those that are also longer in duration, and thus the difference in vowel quality between shorter and longer tokens is sharper for all groups than it is for 1970s Adults. Rather than the correlation between length and peripherality being a straightforward phonetic property of (er), these results suggest that a relationship between the duration of (er) and its position in the vowel space has emerged over time. Separate Pearson's correlations between (er) duration and both F1 and F2 calculated within each age group reveal that all age groups, except for 1970s Adults, show a significant correlation between duration and both F1 and F2, and this relationship gets stronger over time, indicated by the increasing r value (F1: 1970s Adults $r(295) = 0.062$, $p = 0.284$; 1970s Teens: $r(1,157) = 0.304$, $p < 0.001$; 2010s Adults: $r(861) = 0.241$, $p < 0.001$; 2010s Young Adults: $r(1,102) = 0.312$, $p < 0.001$; F2: 1970s Adults $r(295) = -0.038$, $p = 0.518$; 1970s Teens: $r(1,157) = -0.129$, $p < 0.001$; 2010s Adults: $r(861) = -0.166$, $p < 0.001$; 2010s Young Adults: $r(1,102) = -0.227$, $p < 0.001$).

While changes in duration, F1 and F2 are all relevant to the study of (er) over time, for expository purposes, we restrict the subsequent discussion to duration.

[9] Appeal intonation accounts for 3% of the IU-final (er) tokens produced by 1970s Adults, 10% for 1970s Teens, 5% for 2010s Adults and 6% for 2010s Young Adults.

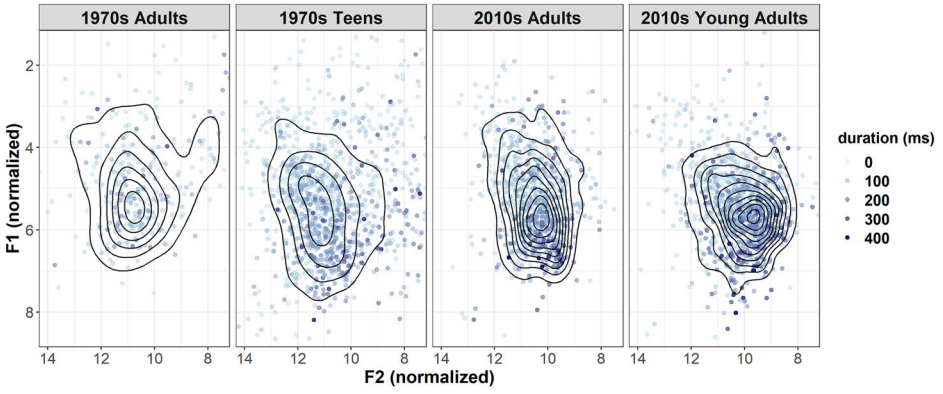

Figure 5 Two-dimensional kernel density plot of normalized F1 and F2 for all speakers over time; longer tokens darker, shorter tokens lighter ($n = 3{,}391$); tokens longer than 400 ms excluded ($n = 32$)

5.3. Lengthening of (er) across the community

As noted above, acoustic work on (er) has identified differentiation across certain ethnic groups, leading to interpretations of (er) as an ethnolectal feature. Based on data from sociolinguistic interviews with 21 speakers recorded in the 1990s, Kiesling proposed that both Greek- and Lebanese-Australians were producing longer and more retracted (er) than Anglo-Australians in 1990s Australia (2005, p. 20), and some 20 years later, based on word list data from 10 speakers, Clothier observed a longer and lower (er) for Lebanese- than Anglo-Australians (2014). The *Sydney Speaks* corpus allows for a real-time test of this proposed ethnolectal feature, providing a large sample of socially stratified participants over different time points.

We first consider in broad terms how each ethnic group has participated in the overall change that we described in §5.1. This is presented in Figure 6, with each of the four age groups now broken down by community, which shows that the lengthening over time does not proceed in the same way across each of the communities. In particular, we observe relative stability in apparent time from the 1970s Anglo Adults to the Anglo Teens (with mean durations of 77 ms and 72 ms respectively). By contrast, both Greek (mean = 110 ms) and to a lesser extent, Italian Teens (mean = 89 ms) in the 1970s produce significantly longer (er) than Anglo Adults and Anglo Teens. While we do not currently have the data available to continue to track (er) realizations for Greek-Australians, we do see further change across the broader community. For the 2010s Adults, both Anglos (mean = 107 ms) and Italians (mean = 116 ms) have lengthened relative to their 1970s counterparts, and both produce (er) with durations similar to those of the 1970s Greek and Italian Teens. For the 2010s Young Adults, while the Anglos have lengthened further (mean = 133 ms), the Italians (mean = 111 ms) retain a similar duration to that of 2010s Italian Adults, and the Chinese-Australians (mean = 107 ms) exhibit an (er) that is longer than earlier realizations, but shorter than that of their 2010s Anglo Young Adult peers.

The patterning for the 1970s Teens is consistent with proposals around the ethnolectal nature of (er). That the 2010s Young Adults produce (er) realizations that resemble those

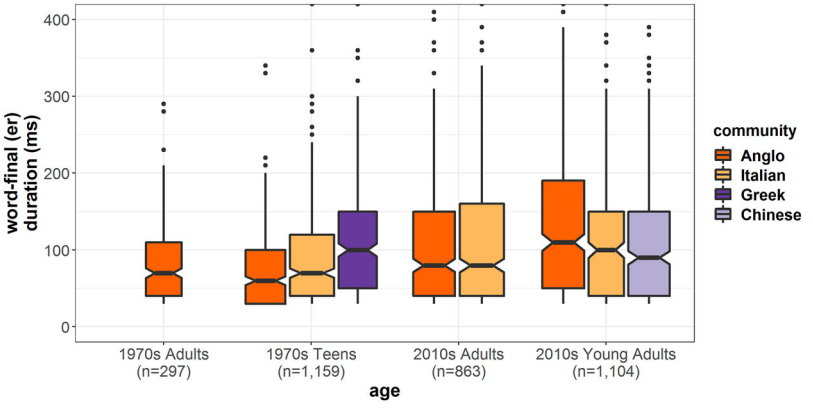

Figure 6 Duration of (er) for all speakers over time, community and sex in IU-final position ($n = 3,423$): Anglo $n = 1,478$; Italian $n = 1,203$; Greek $n = 340$; Chinese $n = 402$

of this earlier time suggests that this ethnic variation has spread more widely, and has now been taken up across the community.

5.4. Ethnicity, class and sex over time

To better understand how this has taken place, we now consider the effects of sex and class alongside community at each time point. The results for the 1970s Adults are presented in Table 3. Here, we see that neither sex nor class significantly impacts the duration of (er); the only significant predictor from the model is speech rate, which, as expected, negatively correlates with (er) duration. At this time point, then, we observe relative homogeneity among Anglo-Australians in their realization of this variable across social groups.

Moving on to 1970s Teens, the model in Table 4 corroborates the lengthening described in §5.1. Compared with 1970s Anglo Teens, both Greek and Italian Teens produce longer (er), with Greeks having the longest realizations by far. Table 4 also shows that, like 1970s Anglo Adults, 1970s Teens are not distinguished socially by sex or class. That (er) lengthening is restricted primarily to Greek- and Italian-Australians at this time supports the interpretation of longer (er) as an ethnolectal feature in the 1970s.

What changes do we see from the 1970s Teens to the 2010s Adults, remembering that this is the same generation, recorded 40 years on? Figure 7 shows the duration of (er) for 2010s Adults by community, sex and class, and Table 5 reports the model fit to the same data. The ethnic variation evident in the 1970s Teens between the Anglos and the Italians

Table 3 Output of linear mixed effects fit to IU-final (er) duration of 1970s Anglo Adults ($n = 297$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=1970s Adult Middle Class Females]	110.49	12.40	8.91	**<0.001**
Male	-9.93	9.95	-1.00	0.318
Lower Working Class	0.67	12.24	0.05	0.956
Upper Working Class	-0.74	11.69	-0.06	0.950
Speech rate	-5.55	1.75	-3.17	**0.002**

Table 4 Output of linear mixed effects fit to IU-final (er) duration of 1970s Anglo, Greek and Italian Teens ($n = 1{,}159$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=1970s Teen Anglo Middle Class Females]	106.75	8.64	12.36	**<0.001**
Male	-8.17	5.96	-1.37	0.171
Lower Working Class	2.93	7.33	0.40	0.689
Upper Working Class	4.48	7.76	0.58	0.563
Greek	36.53	7.52	4.86	**<0.001**
Italian	14.01	7.15	1.96	**0.050**
Speech rate	-6.56	0.99	-6.64	**<0.001**

Note: Anglo $n = 365$; Greek $n = 340$; Italian $n = 454$.

has now been lost, as there is no longer a significant difference between these two groups. An effect for sex, however, emerges, such that females tend to produce longer (er) than males. Furthermore, while males exhibit very little class differentiation, females show clear class stratification. For the Anglo females (on the left), the Lower Working Class produces the longest (er), while for the Italian females (on the right), the Lower and Upper Working Class produce the longest (er). This suggests that, at some point between the 1970s and 2010s, (er) extended from being an ethnolectal feature to, at least in part, a feature of working class female speech, but was resisted by Middle Class Italians, by Middle and Upper Working Class Anglos, and by males. The extension of this ethnolectal feature to, initially, the working class is no doubt attributable to an association between ethnolectal features and low prestige, or non-standard, use (Warren, 1999, p. 91), an observation that has been made more broadly about ethnolectal features (cf. Eckert, 2008, p. 27).

Turning to the youngest population in the sample, the 2010s Young Adults, we observe again that the social factors that played a significant role in the realization of (er) are different from previous age groups. Table 6 shows the results from the model fit to this age group. Because of the distinct class representations across communities

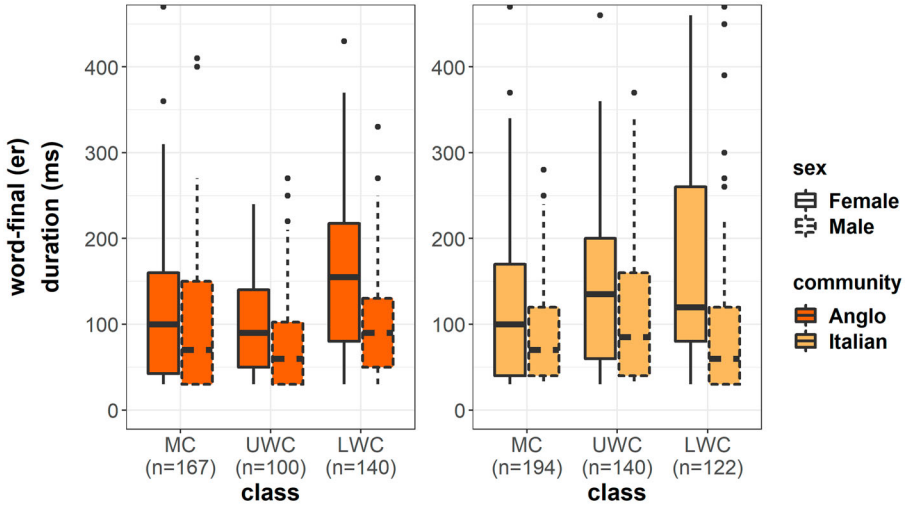

Figure 7 Duration of (er) for 2010s Adults across community, sex and class in IU-final position ($n = 863$): Anglo females $n = 200$; Anglo males $n = 207$; Italian females $n = 209$; Italian males $n = 247$

Table 5 Output of linear mixed effects fit to IU-final (er) duration of 2010s Anglo and Italian Adults ($n = 863$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=2010s Adult Anglo Middle Class Females]	157.72	14.02	11.25	**<0.001**
Male	-32.74	10.90	-2.97	**0.003**
Lower Working Class	28.86	13.54	2.13	**0.033**
Upper Working Class	2.29	12.74	0.18	0.858
Italian	13.74	10.90	1.26	0.208
Speech rate	-9.73	1.77	-5.51	**<0.001**

(specifically the lack of Lower Working Class Chinese-Australians, see Table 1), class was not included in this model. Unlike 2010s Adults, sex does not return significance, indicating that the tendency for longer (er) in female speakers has been lost in this younger generation. In terms of ethnicity, Anglo-Australians exhibit slightly (but not significantly) longer (er) durations than Italian-Australians, and significantly longer durations than Chinese-Australians.

To further explore these apparent differences, separate linear mixed-effects models were fit to each community – Anglo-, Italian- and Chinese-Australian – with class, sex and speech rate as predictors. For the Anglo- and Italian-Australians, neither class nor sex surfaces as significant, indicating that the class distinctions that were evident in the 2010s Adults have been lost for these younger speakers, mirroring an effect also found for Young Anglo Australians for some diphthongs (Grama et al., to appear). We recognize the low numbers of Italian participants in some of the class–sex combinations, but note that the Italian sample is skewed towards the Upper and Lower Working Class, while the reverse is true for the Anglo sample (see Table 1). Thus, if a class effect remained, we might expect the Italians to have longer, not shorter, (er).

However, for the Chinese-Australians, an identical model, depicted in Table 7, reveals a significant interaction between class and sex. As illustrated by the raw values presented in Figure 8, female and male Chinese-Australians show class asymmetry. Middle Class females exhibit significantly shorter (er) than Upper Working Class females, while Middle Class males exhibit longer (er) than Upper Working Class males (though a relevelled model shows that this latter difference is not significant; $\beta = -31.72$, $df = 394$, $p = 0.072$). Put another way, compared with Young Adult Anglo- and Italian-Australians (who have means of 133 and 111 ms respectively), Chinese-Australian Upper Working Class males and Middle Class females exhibit more conservative behaviour – that is, shorter (er) (with means of 82 and 83 ms respectively) – while Upper Working Class Chinese females and Middle Class males (with means of 142 and 125 ms) are similar to Young Adult Anglos and Italians.

Table 6 Output of linear mixed effects fit to IU-final (er) duration of 2010s Anglo, Italian and Chinese Young Adults ($n = 1,104$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=2010s Young Adult Anglo Females]	201.31	11.01	18.29	**<0.001**
Male	-11.37	8.42	-1.35	0.177
Chinese	-19.24	9.67	-1.99	**0.047**
Italian	-19.83	10.74	-1.85	0.065
Speech rate	-13.50	1.55	-8.74	**<0.001**

Note: Anglo $n = 409$; Italian $n = 293$; Chinese $n = 402$.

Table 7 Output of linear mixed effects fit to IU-final (er) duration of 2010s Chinese Young Adults (n = 402)

	Estimate	Std. error	t value	p value
(Intercept) [=2010s Young Adult Chinese Middle Class Females]	140.69	16.37	8.59	**<0.001**
Male	34.41	15.60	2.21	**0.027**
Upper Working Class	55.34	16.67	3.32	**0.001**
Speech rate	-11.15	2.31	-4.82	**<0.001**
Male: Upper Working Class	-87.06	24.31	-3.58	**<0.001**

The profile of the Chinese-Australians is relevant here, in particular their higher socioeconomic status as a community, as described above. Chinese-Australians have been found to pattern similarly to Middle Class Anglo Australians for other phonological variables, and they have also been found to pattern differently for gender, in that males and females equally favour overtly prestigious forms of the diphthongs FLEECE and FACE (Grama et al., to appear). The avoidance of lengthened (er) by this group might be related to its association in the previous generation with the Lower Working Class, despite the lack of such stratification in Young Adult Anglo- and Italian-Australians. However, why this behaviour appears largely relegated to men and women of different class designations remains unclear at present. We note that the diphthongs of 1970s Italian and Greek Teens appear to pattern differently from Anglos with respect to class (Horvath, 1985, p. 79), suggesting that more recent migrant groups may not reproduce the class-based variation of the dominant group.

What these data do clearly indicate is that what was an ethnolectal feature in the 1970s has been adopted across the community, such that (er) has lengthened for all (including the dominant, hegemonic group), though to varying extents. Chinese-Australians in particular do not show homogeneous behaviour in their use of longer (er), being sensitive to class and sex-based variation. Inclusion of contemporary Greek-Australians would shed further light on the nature of (er) across the community today, but this must be left for future work.

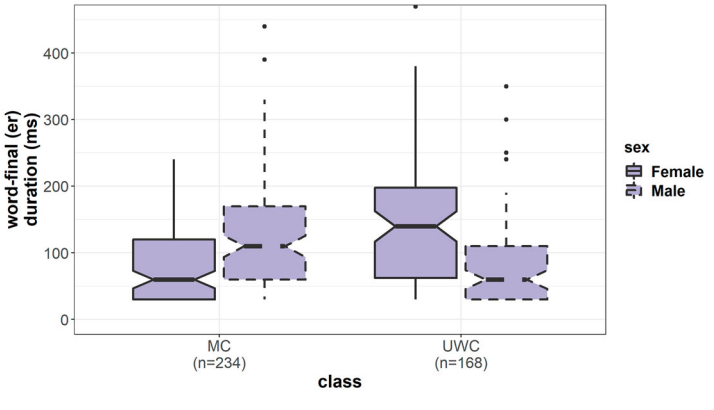

Figure 8 Duration of (er) for 2010s Chinese Young Adults across sex and class in IU-final position: Chinese females n = 205; Chinese males n = 197

6. Conclusion

In tracking (er) over real and apparent time in a large, socially stratified sample, we have been able to make a number of observations about the course of this change. First, we find that (er) has lengthened over time, and that this change has taken place specifically in IU-final position in open syllables. The propagation of the change is therefore linguistically constrained to an environment that is particularly propitious to lengthening. We have further delineated word-final (er) as characterized not just by length, but also by position in the vowel space, and we have shown that the relationship between the position and duration of (er) gets stronger over time.

Second, we identify Greek, and, to a lesser extent, Italian, teenagers in the 1970s as the leaders in the early stages of this change, the very same groups that Horvath (1985) identified as the principal drivers of change in diphthongs towards more 'general' Australian English variants. The change in (er) was subsequently adopted by working class females at some point between the 1970s and 2010s, and today, has been largely adopted across the community by young Australians. The leading role of working class females in taking up this change is consistent with a change from below (Labov, 1990, p. 215); however, the feature is not completely unremarked upon in Australia as either ethnolectal or working class. The 1980s TV show *The Comedy Company* (McMillan et al., 2009) parodied this form in the Greek-Australian character 'Con the fruiterer', in his catch phrase *Doesn't madda*, with a markedly lengthened (er). And some 20 years on, another comedy *Kath & Kim* (Riley & Turner, 2002) parodied it as a working class Anglo feature in the show's opening episode, with Kim declaring "My marriage is over. O-V-A-H!". Although presented as comedy, these caricatures are astute representations of (er)'s changing social associations in Australian English.

We noted above the relationship between ethnolectal features and working class speech (Eckert, 2008, p. 27; Warren, 1999, p. 91). In support of this, early perceptual studies of attitudes to English spoken by ethnic minorities in Australia found that Greek-Australians tended to be evaluated as having lower status than Anglo-Australians, and Italians, as less educated (Callan et al., 1983, p. 417; Gallois & Callan, 1989, p. 155). Thus, it may not be surprising that the adoption of longer (er) by Anglo-Australians was via the working class. That it has been vigorously taken up today by both the dominant hegemonic group (Anglo-Australians) and an established migrant group (Italian-Australians) across the socioeconomic spectrum is evidence of word-final (er)'s strong foothold as a feature of the English of young Australians today. However, the relatively conservative nature of the more recent migrant group (Chinese-Australians) underscores that the feature may retain associations with an ethnic working class. Investigating the behaviour of Greek-Australians and other more recent migrant groups (e.g. Lebanese- and Vietnamese-Australians), as well as interactions between class, sex and ethnicity in these groups, is vital to further understanding this variable and ethnic variation and change more broadly. That features associated with ethnic minorities may be taken up by the wider community, and that other social features, such as class and sex, function alongside ethnicity in conditioning variation, highlights that "there is no obvious way to distinguish between a dialect with ethnic features and an ethnolect" (Eckert, 2008, p. 27).

More generally, these findings support the notion that ethnic minorities can lead in change. There have been several suggestions as to why this might be the case. One proposal is that ethnic minorities have more diverse networks, thus facilitating the spread of innovations (Cheshire et al., 2008). We lack the social network information for 1970s speakers to appropriately capture the moment of the change's inception, though we do observe that the 2010s ethnic minorities have more diverse networks than their Anglo peers. Another proposal is that, rather than drawing on their ethnic background to mark their difference, ethnic minorities draw on changes taking place in the broader society (Gross et al., 2016, p. 245). Again, this is a possible scenario in the data considered here; while long (er) is uncommon in the speech of 1970s Anglos, it is not unattested, so it is conceivable that Greek- and Italian-Australians drew on this feature from the available pool. Along similar lines, it has been noted that ethnic groups may take on features of the majority group to highlight their membership in that group, and differentiate themselves from their accented parents (Callan et al., 1983, p. 410), something which may be particularly relevant for the teenagers recorded in the 1970s (Horvath, 1985, p. 176). We propose a related hypothesis, that second-generation migrants may be proportionally more exposed to, and therefore more likely to adopt, newer features because they have less access to older models of Australian English as their parents are not speakers of that variety. All of these are possible accounts of what we observe, and no doubt there are a combination of factors working together.

What we have seen here is that the status of (er) as an ethnolectal feature in the 1970s shifted to one of working class females at some stage over the past 40 years, and then extended into the general population, especially in the speech of young Anglo- and Italian-Australians. The malleability of word-final (er) demonstrates that ethnolectal features need not be stable over time, but may be adopted by the wider community and that ethnic minorities can lead in language change.

Acknowledgements

We gratefully acknowledge support from the ARC Centre of Excellence for the Dynamics of Language and the *Sydney Speaks* team. We thank Benjamin Purser for his work in establishing the class clusters, and Benjamin Purser, Ksenia Gnevsheva and two anonymous reviewers for valuable feedback on the paper.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by ARC Centre of Excellence for the Dynamics of Language: [Grant Number CE140100041].

Notes on Contributors

James Grama is a Research Fellow and Lecturer in the Sociolinguistics Lab in the Department of Anglophone Studies at the University of Duisburg-Essen. From 2017-2020, he was a Postdoctoral Fellow at the ARC Centre of Excellence for the Dynamics of Language at the Australian National

University, where he worked on the Sydney Speaks project. He holds a Ph.D. in linguistics from the University of Hawai'i at Mānoa. His research focuses on sociophonetic questions of language variation and change, with an emphasis on vowel behaviour. He has worked largely on varieties of English (in e.g. Australia, Hawai'i, and California), English-based creoles (e.g. Hawai'i Creole, Bislama), and minority languages (e.g. Matukar Panau).

Catherine E. Travis is Professor of Modern European Languages in the School of Literature, Languages and Linguistics at the Australian National University, and a Chief Investigator in the ARC Centre of Excellence for the Dynamics of Language. Her research addresses questions related to linguistic and social factors impacting on language variation and change, in particular in socially diverse communities. She leads the Sydney Speaks project on which this paper is based, and has also worked extensively on the study of variation in Spanish and in bilingual speech.

Simon Gonzalez is a Research Officer at the ARC Centre of Excellence for the Dynamics of Language (CoEDL) at the Australian National University (ANU), where he works on the Ku Waru Child Language Socialization Study. From 2017-2020, he was a Postdoctoral Fellow on the Sydney Speaks project and the Transcription Acceleration Project at CoEDL at the ANU. His research focuses on acoustic phonetics, empowered by computational technologies. He develops computational tools (scripts and online apps) for more efficient and practical analysis of phonetic and phonological phenomena. He also makes use of programming languages to facilitate the forced alignment of speech data, in both majority and minority languages.

ORCID

James Grama http://orcid.org/0000-0003-3642-5812
Catherine E. Travis http://orcid.org/0000-0002-1410-3268
Simon Gonzalez http://orcid.org/0000-0001-8013-7821

References

Albin, A. (2014). PraatR: An architecture for controlling the phonetics software "Praat" with the R programming language. *The Journal of the Acoustical Society of America*, *135*(4), 2198–2199.
Australian Bureau of Statistics. (1981). Census of population and housing.
Australian Bureau of Statistics. (1986). Census of population and housing.
Australian Bureau of Statistics. (1991). Census of population and housing.
Australian Bureau of Statistics. (2016a). Ancestry fact sheet.
Australian Bureau of Statistics. (2016b). Census of population and housing.
Australian Curriculum Assessment and Reporting Authority. (2019). My school. https://www.myschool.edu.au/
Australian Qualifications Framework Council. (2013). Australian qualifications framework January 2013. (Vol. 2019).
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01.
Bladon, R. A. W., Henton, C. G., & Pickering, J. B. (1984). Towards an auditory theory of speaker normalization. *Language and Communication*, *4*, 59–69.
Bořil, T., & Skarnitzl, R. (2016). Tools rPraat and mPraat. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Text, Speech and Dialogue* (pp. 367–374). Cham(Switzerland): Springer International Publishing.
Callan, V. J., Gallois, C., & Forbes, P. A. (1983). Evaluative reactions to accented English: Ethnicity, sex role and context. *Journal of Cross-Cultural Psychology*, *14*(4), 407–426.
Chafe, W. (1994). *Discourse, consciousness and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, IL: University of Chicago Press.
Cheshire, J., Fox, S., Kerswill, P., & Torgersen, E. (2008). Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica*, *22*(1), 1–23.

Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: the emergence of multicultural London English. *Journal of Sociolinguistics*, *15*(2), 151–196.

Clothier, J. 2014. *Heading South: Phonetic differences in word finaler in speakers from Lebanese- and Anglo-Celtic Australian ethnic groups*. Paper presented at the Australian Linguistic Society Annual Conference, University of Newcastle, 10–12 December 2014.

Clothier, J. (2019). A sociophonetic analysis of /l/ darkness and Lebanese Australian ethnic identity in Australian English. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences (ICPhS)* (pp. 1888–1892). Melbourne, Australia: Australasian Speech Science and Technology Association Inc.

Clothier, J. (2020). Ethnolectal variability in Australian Englishes. In L. Willoughby & H. Manns (Eds.), *Australian English reimagined: Structure, features and developments* (pp. 155–171). New York, NY: Routledge.

Clothier, J., & Loakes, D. (2018). Coronal stop VOT in Australian English: Lebanese Australians and mainstream Australian English speakers. In J. Epps, J. Wolfe, J. Smith, & C. Jones (Eds.), *Proceedings of the 17th Australasian International Conference on Speech Science and Technology (SST)* (pp. 13–16). Sydney: Australasian Speech Science and Technology Association Inc.

Clyne, M., Eisikovits, E., & Tollfree, L. (2001). Ethnic varieties of Australian English. In D. Blair & P. Collins (Eds.), *English in Australia* (pp. 223–238). Amsterdam: John Benjamins.

Clyne, M., Eisikovits, E., & Tollfree, L. (2002). Ethnolects as in-group varieties. In A. Duszak (Ed.), *Us and others: Social identities across languages, discourses and cultures* (pp. 133–157). Philadelphia, PA: John Benjamins.

Clyne, M., & Kipp, S. (1997). Trends and changes in home language use and shift in Australia, 1986–1996. *Journal of Multilingual and Multicultural Development*, *18*(6), 451–473.

Cox, F., & Palethorpe, S. (2011). Timing differences in the VC rhythm of standard Australian English and Lebanese Australian English. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)* (pp. 528–531). Hong Kong: City University of Hong Kong.

Cox, F., & Palethorpe, S. (2012). Standard Australian English: The sociostylistic broadness continuum. In R. Hickey (Ed.), *Standards of English: Codified varieties around the world* (pp. 294–317). Cambridge: Cambridge University Press.

Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. Edwards & M. Lampert (Eds.), *Talking data: Transcription and coding in discourse* (pp. 45–89). Hillsdale: Lawrence Erlbaum Associates.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism*, *12*(1&2), 25–42.

Fletcher, J., & McVeigh, A. (1993). Segment and syllable duration in Australian English. *Speech Communication*, *13*, 355–365.

Flynn, N., & Foulkes, P. (2011). Comparing vowel formant normalization methods. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)* (pp. 683–686). Hong Kong: City University of Hong Kong.

Fromont, R., & Hay, J. (2012). LaBB-CAT: An annotation store. In *Proceedings of the Australasian Language Technology Workshop* (pp. 113–117).

Gallois, C., & Callan, V. J. (1989). Attitudes to spoken Australian English: Judgements of ingroup and ethnic outgroup speakers. *Australian Journal of Linguistics*, *9*(1), 149–160.

Gonzalez, S., Grama, J., & Travis, C. E. (2020). Comparing the performance of forced aligners used in sociophonetic research. *Linguistics Vanguard*, *6*(1). doi:10.1515/lingvan-2019-0058.

Grama, J., Travis, C. E., & Gonzalez, S. To Appear. Ethnic variation in real time: Change in Australian English diphthongs. In H. Van de Velde, N. Haug Hilton, & R. Knooihuizen (Eds.), *Studies in Language Variation, Papers from the tenth International Conference on Language Variation in Europe (ICLaVE 10)*, Leeuwarden, June 2019), Amsterdam: John Benjamins.

Gross, J., Boyd, S., Leinonen, T., & Walker, J. A. (2016). A tale of two cities (and one vowel): Sociolinguistic variation in Swedish. *Language Variation and Change*, *28*(2), 225–247.

Guy, G. R., Horvath, B. M., Vonwiller, J., Daisley, E., & Rogers, I. (1986). An intonational change in progress in Australian English. *Language in Society*, *15*(1), 23–51.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, *28*, 100–108. doi:10.2307/2346830.
Hoffman, M. F., & Walker, J. A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, *22*(1), 37–67.
Horvath, B. (1985). *Variation in Australian English: The sociolects of Sydney*. Cambridge: Cambridge University Press.
Jupp, J. (2001). *The Australian people: An encyclopedia of the nation, its people and their origins* (2nd ed.). Cambridge: Cambridge University Press.
Kassambara, A., & Mundt, F. (2020). factoextra: Extract and visualize the results of multivariate data analyses. R package 1.0.7. https://cran.r-project.org/package=factoextra.
Kiesling, S. F. (2005). Variation, stance and style: Word-final *-er*, high rising tone, and ethnicity in Australian English. *English World-Wide*, *26*(1), 1–42.
Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America*, *59*(5), 1208–1212.
Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.
Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular*. Philadelphia, PA: University of Pennsylvania Press.
Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28–53). Englewood Cliffs, NJ: Prentice Hall.
Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, *2*(2), 205–254.
Labov, W. (2008). Mysteries of the substrate. In M. Meyerhoff & N. Nagy (Eds.), *Social lives in language – Sociolinguistics and multilingual speech communities* (pp. 315–326). Amsterdam/Philadelphia: John Benjamins.
Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, Instruments, & Computers (Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. http://tla.mpi.nl/tools/tla-tools/elan/)*, *41*(3), 841–849.
Liljencrants, J., & Lindblom, B. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. *Language*, *48*(4), 839–862.
Lüdecke, D., & Schwemmer, C. (2020). sjPlot: Data visualization for statistics in social science. R package version 2.8.2. https://CRAN.R-project.org/package=sjPlot.
McMillan, J., Beavis, A., & Jones, F. L. (2009). The AUSEI06: A new socioeconomic index for Australia. *Journal of Sociology*, *45*(2), 123–149.
Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge: MIT Press.
R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org.
Riley, G., & Turner, J. (2002). Kath and Kim. T Emery (Director), Sex. Melbourne, Australia: ABC TV.
Sala, E., Dandy, J., & Rapley, M. (2010). 'Real Italians and Wogs': The discursive construction of Italian identity among first generation Italian immigrants in Western Australia. *Journal of Community & Applied Social Psychology*, *20*, 110–124. doi:10.1002/casp.1023.
Travis, C. E. (2016–2021). *Sydney speaks*. Australian Research Council Centre of Excellence for the Dynamics of Language, Australian National University. http://www.dynamicsoflanguage.edu.au/sydney-speaks/.
Travis, C. E., Grama, J., & Gonzalez, S. (In Progress). *Sydney Speaks corpora*. Australian Research Council Centre of Excellence for the Dynamics of Language, Australian National University. http://www.dynamicsoflanguage.edu.au/sydney-speaks/.
Warren, J. (1999). Wogspeak: The transformation of English in Melbourne. *Journal of Australian Studies*, *62*, 86–94.

Willoughby, L., & Manns, H. (2020). Introducing Australian English. In L. Willoughby & H. Manns (Eds.), *Australian English reimagined: Structure, features and developments* (pp. 1–12). New York (NY): Routledge.

Appendix

Table A1 Output of linear mixed effects fit to (er) duration of IU-final contexts, with age and transitional continuity (TC) modelled as main effects and an interaction, and morphology, and speech rate as predictors ($n = 3,423$)

	Estimate	Std. error	*t* value	*p* value
(Intercept) [=1970s Adults, TC final, monomorphemic]	110.62	9.91	11.16	**<0.001**
1970s Teens	2.23	10.08	0.22	0.825
2010s Adults	29.63	10.49	2.82	**0.005**
2010s Young Adults	37.06	10.15	3.65	**<0.001**
TC continuing	19.27	8.46	2.28	**0.023**
TC appeal	17.62	24.55	0.72	0.473
Polymorphemic	1.89	4.68	0.40	0.686
Speech rate	-8.86	0.76	-11.69	**<0.001**
1970s Teens: TC continuing	0.75	9.62	0.08	0.938
2010s Adults: TC continuing	5.06	9.84	0.51	0.607
2010s Young Adults: TC continuing	-0.97	9.57	-0.10	0.919
1970s Teens: TC appeal	29.53	25.73	1.15	0.251
2010s Adults: TC appeal	15.67	27.10	0.58	0.563
2010s Young Adults: TC appeal	17.77	26.29	0.68	0.499