English Language and Linguistics 17.2: 359–390. © Cambridge University Press 2013
doi:10.1017/S1360674313000063

'Flip-flop' and mergers-in-progress[1]

LAUREN HALL-LEW
University of Edinburgh
(Received 27 April 2012; revised 16 January 2013)

During a merger-in-progress, occasionally one or two speakers will exhibit an unusual phonological pattern reminiscent of *flip-flop* (Labov *et al.* 1972). In such cases, the merging vowels appear to move past the point of coalescence in at least one phonetic dimension; difference is maintained but the vowel quality is opposite to the historical pattern on one or both dimensions. Flip-flop between the COT and CAUGHT vowels occurs for two speakers in a recent sample from San Francisco, California. The community shows robust change in progress toward a lower and fronter CAUGHT vowel nucleus, and no change in apparent time for COT. Further analysis shows that this is leading to a change in apparent time toward merger, and that the rate of vowel convergence is stronger among Chinese Americans than European Americans. The two speakers who produce flip-flop are seen to represent a key transitional generation with respect to the ethnic identity of the neighborhood, where flip-flop may be but one linguistic consequence of a lifetime of active negotiation between conflicting local meanings. The analysis suggests that ethnographic detail and attention to individual outliers allows for more comprehensive models of the range of phenomena associated with vowel mergers.

1 Introduction

In a community undergoing merger-in-progress, occasionally one or two speakers will exhibit an unusual phonological pattern reminiscent of *flip-flop* (Labov *et al.* 1972; Labov 1994: 143; Di Paolo 1992: 281). In such cases, the merging vowels appear to move past the point of coalescence in at least one phonetic dimension; difference is maintained but the vowel quality is opposite to the historical pattern on one or both dimensions. The present article looks at flip-flop in a community undergoing COT–CAUGHT merger in San Francisco, California in 2008–9. The data show a robust change in progress toward a lower and fronter CAUGHT vowel and no change in apparent time for COT; for the youngest speakers this results in complete overlap between vowel classes. Two of the thirty individuals analyzed are seen producing a flip-flop pattern, with one speaker producing CAUGHT lower than COT and another speaker producing CAUGHT further front than COT. Drawing on the diagnostic value of individual random intercepts

[1] The first version of this work was presented at the International Society for the Linguistics of English (ISLE), in June 2011, on the panel 'Mergers in English: perspectives from phonology, sociolinguistics and psycholinguistics', organized by Lynn Clark, Warren Maguire and Kevin Watson. I extend my thanks to them for organizing the workshop and this special issue. A second version was presented at the Department of Linguistics at Queen Mary, University of London, in January 2012, and I thank that audience for their valuable feedback. I am indebted to the anonymous reviewers of this paper for their detailed and insightful observations. Thanks to Jennifer Nycz, Kevin Stadler, Helen West and Amy Wong for useful discussions. The speakers whose voices are analyzed here receive my greatest debt of gratitude.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

(Drager & Hay 2012), these speakers clearly appear as statistical outliers in the sample. Further comparison between interview and wordlist styles reveals that both speakers dramatically reduce vowel distinction in the wordlist condition, suggesting an analysis of flip-flop as somehow related to near-merger. Although tentative, ethnographic analysis points to a set of potential social motivations for flip-flop.

Although flip-flop is rare, it is not unprecedented. Several studies of the North American COT–CAUGHT merger have had at least one speaker who flip-flops. Labov (1994: 355) notes that Herold (1997) removed one speaker from the analysis 'because the speaker pronounced *Don* as [dɔn] and *dawn* as [dɑn]'. One token may be a speech error, but other studies find more regular flip-flopping. In Utah, Di Paolo (1992) shows 'Speaker 251' flip-flopping along two of the three phonetic dimensions analyzed (F1 and the Voice Quality Index, but not F2). Boberg (personal communication) reports the occurrence of COT–CAUGHT flip-flop among some speakers of Canadian English (see Boberg 2010), and Nycz (personal communication) finds the same result for one Canadian migrant to New York City (see Nycz 2011).

Flip-flop is a theoretically surprising production pattern, falling outside the cline of more expected configurations for merger-in-progress: distinction, merger and near-merger. Near-merger (Labov *et al.* 1972; Labov *et al.* 1991; Yu 2007) essentially refers to speakers who produce a distinction but perceive a merger. Typically, the phonetic difference between the distinct phonemes is small ('less than 200 Hz', Labov *et al.* 2006: 285). When a community exhibiting near-merger includes (older) speakers with complete distinction and (younger) speakers with complete merger, the near-merged speakers can be thought of as sharing the underlying phonological inventory of the older speakers, but with acoustic targets closer to those produced by the younger speakers (see also Nycz 2005 and Yu 2007 for phonological accounts of near-mergers). Theoretically, the progress from distinction to merger proceeds from one well-defined point to another, with varying degrees of near-merger representing the variation between those points; Labov (2007) notes that 'near-mergers give us a static view' of how mergers progress through a community. The phenomenon of flip-flop, however, appears to lie outside this range of variation, because rather than the acoustic difference between the two phonemes growing increasingly smaller, the distinction remains but 'flips' along a particular phonetic dimension. Attempting to understand what flip-flop is and why it might occur may enrich our understanding of how mergers actually progress through a community.

For flip-flop to occur a speaker must have a phonemic distinction; at one level, 'flip-floppers' are just speakers with either distinction or near-merger, the difference being whether they perceive the distinction they make or not. What makes them distinct is that they have acquired an additional motivation to aim for unexpected phonetic targets. There are two questions, then: first, do flip-flopped speakers perceive their phonological distinction, and second, why is their phonetic target different? Evidence (see section 4) suggests that the answer to the first question is no, and speakers who flip-flop are exhibiting a special kind of near-merger. The answer to the second question relies on modeling the relevant phonetic targets for low back merger in the community. Crucially,

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

the merger is accomplished by a shift in CAUGHT, not COT, and so flip-flop most likely results from a speaker aiming for an advanced production of CAUGHT that 'overshoots' even the most advanced realizations in the community. Ethnographic evidence (see section 5) may help account for why this might occur.

1.1 COT *and* CAUGHT *in San Francisco English*

The low back vowels have been through a number of changes throughout the history of English, resulting today in a complex distribution of lexical sets that variably correspond to different phonemic classes in different areas of the English-speaking world. The low back merger in North American English results from the loss of a distinction between the 'short-*o*' class of COT and the 'long-*o*' class of CAUGHT (Labov *et al.* 2006, henceforth ANAE). Labov (1994: 316) calls the COT–CAUGHT merger 'the largest single phonological change taking place in American English'. It has been analyzed extensively in dialectology (Kurath & McDavid 1961; ANAE), with many studies showing its rapid spread across North America (Labov *et al.* 1972; Herold 1997; Majors 2005; Gordon 2006; Eberhardt 2008). Several studies in the Western states have argued for the presence of merger (e.g. Terrell 1975 (cited in Labov 2010); Moonwomon 1991; Fridland 2004; though see Di Paolo 1992). Labov (1998) suggests that speakers in the West and Canada together form a broadly unified dialect area by virtue of their shared production of the merger, which puts them in contrast to North Americans who are in phases of transition or who maintain the distinction.

Interestingly, while the ANAE documents merger across the Western US, San Francisco is a singular exception, characterized as 'transitional' (ANAE: 61). The only Western US speaker in the ANAE with complete distinction is a San Franciscan, and the other five San Franciscans are classified as transitional, rather than merged. As a result, the ANAE's low back merger isogloss juts around San Francisco, encircling the rest of California while leaving the city an isolated island of distinction. I suggest here that the reasons for this may be rooted in a particular San Franciscan identity that emphasized similarities with cities on the East Coast rather than the Western US (see also Hall-Lew under review). At the same time, the present data support previous assertions (e.g. Moonwomon 1991) that San Francisco is moving toward merger in apparent time.

Indeed, despite its exceptionality, variation in the low back vowels, particularly CAUGHT, is among the oldest attested features of San Francisco English. DeCamp found that '/ɔ/ has a wider phonetic range than any other phoneme' (1959: 59), in particular a 'peculiar use of fronted allophones of /ɔ/' (1953: 555), whereas /ɑ/ showed little variation. Drawing on evidence of merger in other parts of the Western US ('parts of Washington' and 'Utah, for example', DeCamp 1959: 60), he interpreted this variability as evidence 'that this coalescence is beginning in San Francisco' (1953: 555). Other accounts of Western US English (Pederson 2001), and San Francisco in particular (Moonwomon 1991), have similarly characterized the change as a movement of /ɔ/ (CAUGHT) rather than movement of /ɑ/ (COT).

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Low back vowel production has been analyzed in all known studies of phonetic variation in San Francisco English. The results from Hinton *et al.* (1987) foreshadow those of the ANAE, finding that 'the merger is not complete, and while movement of /ɔ/ toward /ɑ/ is indicated, it is not especially vigorous', concluding that 'phonetically, [ɔ] is alive and well in California' (123). In contrast, Moonwomon (1991) argued instead that the merger was actually 'nearly complete' (213), with all of the youngest speakers in her study showing 'complete or almost complete' (203) overlap between /ɑ/ and /ɔ/. The relevant ANAE data are based on six San Franciscans surveyed in 1995. The six speakers (ages 39–72) were all older than the most merged speakers in Moonwomon's (1991) study, and given DeCamp's evidence that the change was only just 'beginning' in the 1950s, the ANAE's finding of San Francisco's apparent time lag may be at least partially due to the lack of participants who were born after the mid 1950s.

The rate of change and the extent to which merger has gone to completion are both open areas of study for San Francisco English. Additionally, the ethnic demographics of the city have shifted dramatically since the 1950s. Earlier studies primarily focused on European Americans, and the present article considers Chinese Americans as well. One of the most salient social changes in the neighborhood under study has been a major ethnic shift from a European American majority to a Chinese American plurality, and in fact this social change will be seen to be important for modeling the contemporary linguistic pattern.

1.2 Social change in the Sunset District

San Francisco is a geographically small city (46.7 square miles), yet very socially diverse. The social landscape is divided into approximately forty neighborhoods, some strongly associated with particular ethnic identities (Godfrey 1988). The Sunset District is the city's largest residential neighborhood, with eighteen census tracts totaling 93,073 residents in 2010 (Bloch *et al.* 2011). Ten of these tracts had grown since 2000; all saw increases in their Asian population and all but one were over 50 percent Asian in 2010. The 2010 population was 51 percent Asian (47,604) and 38 percent white (35,170). Subgroup data were not available for 2010 at the time of writing, but in 2000, 76 percent of these Asian residents were of Chinese descent. At the time of fieldwork in 2008–9, the Sunset had gained a reputation as a 'New Chinatown' (see Laguerre 2005). When residents talked about local social change, the responses overwhelmingly concerned this ethnic shift.

Originally, the Sunset was essentially a collection of Irish Catholic parishes. Because of its late development (Ungaretti 2012), many of the first Sunset residents came from other parts of San Francisco, and its first reputation was as a neighborhood of 'long-time San Franciscans who have moved from the older sections of the city' (DeCamp 1953: 23). This foundation led to a local narrative of authenticity, in which the Sunset became the neighborhood where 'real native San Franciscans' lived. Many of these 'authentic' San Franciscans had come from the city's Mission District, known at that time for its Irish Catholic identity (Hall-Lew under review). As the Sunset changed from an Irish

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

parish to a New Chinatown, its narrative of authenticity appears to have persevered. While local authenticity remains central to local identity, the social meanings of that authenticity have become layered and complex: in general, older speakers recognize Irish Catholic social practices as indexing local identity, while younger speakers associate the Sunset District with Chinese American norms.[2] The result is that two rather different sets of semiotic resources both allow for the construction of native identity, resulting in two different yet concurrent kinds of authenticity. This complex process is possible because authenticity, or more precisely *authentication* (Bucholtz 2003), is a social project and not a static quality (Eckert 2003).

Elsewhere (Hall-Lew forthcoming), I have argued that contemporary Sunset District authenticity is linked to Chinese American ethnic identity and social practice. The process of social change leading to this state is important for the present analysis, for two reasons. First, the 'Irish' authenticity and the 'Chinese' authenticity map on to speech styles: older residents associate San Franciscan identity more with cosmopolitan centers on the East Coast – New York and Boston – than with urban centers in the Western US, and describe their English variety similarly (e.g. 'sort of Brooklyn'; see Hall-Lew under review). Younger residents, in contrast, associate San Franciscan identity more with the rest of Northern California. As a result, different realizations of CAUGHT may index different authenticities (Wong & Hall-Lew under review), and the local valuing of a high, back CAUGHT vowel among older Irish Americans may be one factor inhibiting completion of merger (see the ANAE results). Second, while most individuals orient to either one or the other representation of authenticity, some individuals are caught between the two, with a certain level of access to both sets of meanings, and with the challenging task of negotiating between them. These pressures may be related to those speakers' propensity to exhibit the flip-flop vowel pattern.

2 Methodology

2.1 Data collection

Data analysis is based on 30 speakers (table 1)[3] assembled based on year of birth, sex class and ethnicity. The ethnic contrast is limited to the most salient community contrast: Chinese American ('ChA', N = 16) and European American ('EA', N = 14). The ChA group is mostly Cantonese (N = 13), while EA is much more heterogeneous,

[2] While contemporary Sunset residents describe their population shift by drawing a binary distinction between 'Asian' and 'white', their frequent use of the descriptor 'Chinatown' points to the ethnographic observation that the referent of 'Asian' is often actually 'Chinese'. 'Asian' is subject to a local a recursive process (Irvine & Gal 2000) where 'Asianness' is ideologically ascribed to East Asian rather than non-East Asian, and then to Chinese rather than non-Chinese, and even to Cantonese/Toisanese rather than other Chinese. To avoid confusion here, I focus on the group typically being referred to (Chinese) rather than the local name used for that group (Asian).

[3] The present sample differs from that in Hall-Lew (2009). One speaker ('Sam') excluded from that study due to lack of tokens is included here because tokens of Wells's (1982) CLOTH class are included here. Furthermore, five speakers here replace five analyzed in Hall-Lew (2009) to better balance the set with respect to year of birth, sex class and ethnicity.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Table 1. *Speakers and their classification with respect to social factors*

Pseudonym	Sex class	YOB	Ethnicity
Enid	F	1932	ChA
Jenny	F	1949	ChA
Ruth	F	1954	ChA
Cindy	F	1966	ChA
Emily	F	1970	ChA
Molly	F	1971	ChA
JoJo	F	1982	ChA
Vicky	F	1985	ChA
Monica	F	1991	ChA
Lou	M	1922	ChA
Mickey	M	1945	ChA
Sal	M	1962	ChA
John	M	1977	ChA
Hector	M	1981	ChA
Pete	M	1984	ChA
Skylar	M	1991	ChA
Cheri	F	1942	EA
Ethel	F	1959	EA
Amy	F	1968	EA
Abby	F	1977	EA
Mary	F	1978	EA
April	F	1989	EA
Grace	F	1991	EA
Don	M	1941	EA
Bruce	M	1944	EA
Danny	M	1963	EA
Richard	M	1967	EA
Aaron	M	1979	EA
George	M	1990	EA
Sam	M	1991	EA

the largest subgroup being Irish (N = 5). Demographic balance was sought with respect to all factors but is not exact; the lack of European American speakers born prior to 1940, for example, will be addressed in section 3.

The sample represents neighborhood demographics while still limiting the study to English dominant speakers; whereas 58 percent of Sunset homes are not English dominant, all speakers in this sample are. The non-native English speakers have been English dominant since at least age 5, with the exception of the oldest speaker, 'Lou', who emigrated from China at age 14 (he was 87 when interviewed). Lou is the only one not born in San Francisco; 26 were raised from infancy in the Sunset District and the remaining 3 are older Chinese Americans who spent early childhood in the Chinatown

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

area before changes in fair housing legislation allowed them to move as older children or young adults.

All vowel tokens come from interview[4] or wordlist[5] speech. Only syllables with primary stress were included. Formant values were automatically extracted from hand-labeled TextGrids in *Praat* (Boersma & Weenink 2012), using *Akustyk* (Plichta 2006). TextGrids segmentation boundaries were placed roughly equidistant from the midpoint of the nucleus, which was defined as the point with the highest F1 value (Labov *et al.* 1972; Thomas 2011). End boundaries were placed roughly two glottal pulses from the end of regular voicing or, for pre-nasal tokens, at the point of a drop in amplitude from the oral to the nasal segment.[6] Because of errors in automatic extraction, 18 percent of the total set of tokens were either checked or re-measured by hand.[7] The analysis of interview speech is based only on formant measures, whereas the analysis of wordlist speech includes both formant and duration values; measuring duration for spontaneous speech is complicated by the need to normalize for rate of speech, as well as various discourse factors (Di Paolo *et al.* 2011). Measures of phonation difference (see Di Paolo 1992), and other cues to vowel quality, are left for future analysis.

Tokens of the low back vowels were classified as belonging to either the COT class or the CAUGHT class.[8] Outside sources (ANAE; DeCamp 1953; Wells 1982; Moonwomon 1991; Johnson 2007) were consulted to determine class membership. Words not listed in those sources were assigned based on the consensus intuitions of several sociophoneticians with experience working on the low back vowel merger. Most challenging were instances of Wells's CLOTH class in a voiced velar context, particularly instances of *fog*. DeCamp (1953) found *fog* to be the most fronted '/ɔ/' word. While most US varieties with distinction produce CLOTH /ɔ/, many use /ɑ/ for the words *fog* and *log* (but not *dog* or *long*; Wells 1982). Therefore, in contrast to DeCamp (1953), in the present data set *dog* was coded as CAUGHT and *fog* as COT. Future analysis might analyze these words specifically, but there were too few instances per speaker here.[9]

For interviews, the coding goal was at least ten tokens per vowel class, per speaker. This was often exceeded but not always met, in part because the definition of the envelope of variation. Wells's (1982) PALM lexical set was excluded: *father*, *ma* and *pa* (as well as *mom* and *mommy*) and all 'foreign (a)' loanwords (Boberg 1999, 2009). Following liquids were also excluded, because liquids cause both vowels to

[4] All the interviews were conducted by the author, a female of both Chinese and European descent with family connections to the neighborhood, with merged COT and CAUGHT vowels.

[5] Thanks to Rachel Cristy for assistance with segmenting the wordlist files.

[6] The same models were generated for the vowel off-glide data, based on measurements of all the same tokens. Despite the impressionistic observation that some (older) speakers occasionally produce highly diphthongal, up-gliding CAUGHT vowels, all the models for off-glide data mirrored the models for the nuclei data, exactly as reported in this section. This article thus focuses on data from the vowel nuclei only.

[7] A token was checked if its F1 or F2 standard deviation (across 10ms intervals) was more than 100 Hz.

[8] The keywords COT and CAUGHT are used rather than Wells's (1982) LOT and THOUGHT; COT and CAUGHT include LOT and THOUGHT, respectively, but both also include members of the CLOTH class.

[9] Some San Franciscans will find this surprising, but many interviewees never mentioned fog at all; others noted that the Sunset District is not as foggy 'these days' as it once was. Perhaps climate change has influenced the lexical representation of this token set!

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

back in a process distinct from COT–CAUGHT merger (Moonwomon 1991). The word *aunt* was excluded because of its variable pronunciation as /æ/, and all tokens of *not* were excluded due to their unique potential for prosodic reduction in non-adversarial interactional styles (Yaeger-Dror *et al.* 2003).

The wordlist was comprised of eighty minimal, semi-minimal and homonymic pairs. Five are COT–CAUGHT pairs, four minimal (*dawn–Don*, *cot–caught*, *stalk–stock*, *tot–taught*) and one semi-minimal pair (*dot–daughter*). Five more have COT or CAUGHT paired with a non-low back vowel (*song–sang*, *cost–cast*, *calf–cough*, *bought–boat*, *pot–pout*). All ten were dispersed semi-randomly with respect to sixty minimal pairs (e.g. *well–will*), seven semi-minimal pairs (e.g. *dune–condone*) and three homonymic pairs (or two if *root–route* is excluded; it was variably treated as homonymic or minimal).

All data were normalized using the modified Watt and Fabricius method (Fabricius *et al.* 2009), as realized in the NORM vowel normalization suite (Thomas & Kendall 2007). A speaker's vowel space was represented by COT, CAUGHT, BEET (/i/ in all environments) and COOL (/u/ before /l/, the environment that inhibits fronting in San Francisco English; see Hall-Lew 2011). Interview data and wordlist data were normalized separately.

2.2 Statistical models

The results for the interview speech are based on separate mixed model regressions for F1 and F2, run with the *languageR* package (Baayen 2010), in *R* (R Development Core Team 2010). The independent linguistic factors included vowel identity (VOWEL) and following phonological environment (FOLLOWING). VOWEL refers to COT (N = 469) versus CAUGHT (N = 469). FOLLOWING includes fricatives (regardless of place and voicing), nasals (regardless of place), and alveolar, velar and bilabial stops (regardless of voicing). Open syllables were also included, although COT cannot occur in an open syllable. Since there are no occurrences in this data set of CAUGHT preceding a bilabial, these two environments categorically predict vowel class membership, and are included here to best represent the range of distribution (minus following liquids). Since the data are based on spontaneous speech, levels of FOLLOWING differ in representation, from alveolar stops (N = 352) and velar stops (N = 267) to bilabial stops (N = 133), fricatives (N = 83), nasals (N = 72) and open syllables (N = 31).

The three independent social factors are speaker sex class, year of birth and ethnicity (table 1). Because of differences in interview content and length of interview, the number of tokens per speaker differs, with the most for 'George' (N = 54) and the least for 'April' (N = 17), but no significant correlation was found between the number of tokens per speaker and any of the factors of interest here.

SPEAKER, WORD and SPEAKER:VOWEL were included as random effects. The interaction effect reflects the fact that the data are drawn from a population that includes speakers with complete vowel distinction as well as speakers with complete vowel overlap. Including interaction effects as random factors generates separate random

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

intercepts for each vowel class for each speaker (see Drager & Hay 2012), which is ideal when rates of variability differ between the two vowel classes.

3 Results

The full details of the best models for both interview speech and wordlist speech are presented in the Appendix.

3.1 Interview data – F1 and F2

The only main effects predicting F1[10] variation in the interview data are vowel identity and following phonological environment; none of the social factors are significant main effects. As predicted for any US English variety without complete merger, CAUGHT is higher than COT, overall ($t = -2.913$, indicating a significant disfavoring of a high F1). With respect to phonological environment, a following nasal promotes a higher vowel than a following alveolar stop. This effect is very strong, even more significant than the effect of a following bilabial or an open syllable environment, despite those occurring categorically. The strong effect of a following nasal is perhaps not surprising given that California English also exhibits a nasal system for the BAT vowel, where following nasals promote raising (Eckert 2008b).

One social factor does reach significance as an interaction effect with vowel class. The interaction between year of birth and vowel ($t = 2.877$) indicates that younger speakers favor lower CAUGHT, and that the pattern does not obtain for COT (figure 1). Since CAUGHT is higher than COT overall, it appears to be lowering in the direction of COT. Note that year of birth is not a significant main effect, which contrasts with representations of a merged COT/CAUGHT raising in apparent time in Northern California English (e.g. Eckert 2008b: 34, figure 1). Instead, the nucleus of CAUGHT is lowering in apparent time, and the nucleus of COT is not correlated with age.

The results for F2 are similar, although with different details for following phonological environment. As with F1, the only significant main effects are vowel identity and following phonological environment. CAUGHT is backer than COT ($t = -4.851$), and following nasals are again significant, although this time they predict relatively backer vowels. Following fricatives and velars promote the same effect. Year of birth is again significant as an interaction effect with vowel class ($t = 4.813$), indicating that younger speakers favor fronter CAUGHT vowels, with no pattern for COT (figure 2).

The results show that the first and second midpoint formant frequencies differentiate COT and CAUGHT; low back vowel merger was therefore not a completed sound change in 2008–9. However, the interaction effects do suggest a change in apparent time, where CAUGHT is lowering and fronting, suggesting eventual convergence with COT.

[10] Recall that all references to 'F1' and 'F2' actually refer to normalized values (F1/S(F1) and F2/S(F2), respectively).

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

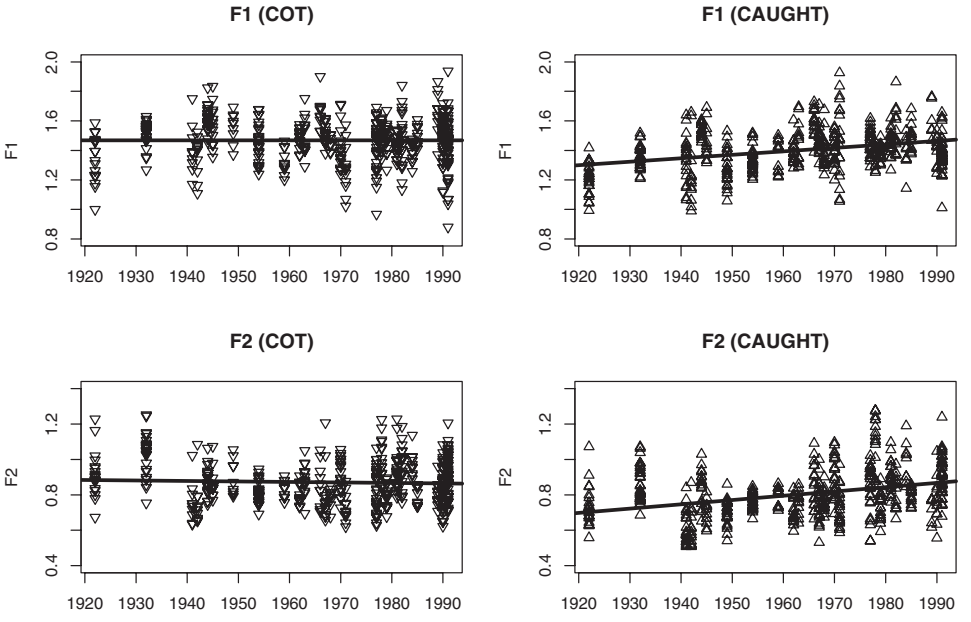


Figure 1. Speaker year of birth with respect to F1 and F2 for COT and CAUGHT

3.2 Interview data: merger

The models predicting F1 and F2 variation do not directly model merger. Since the data are not balanced with respect to following phonological environment, to represent the distance between COT and CAUGHT within each individual speaker's system, an aggregate measure of distance is used: the Euclidean distance between mean values of vowel nuclei (e.g. Irons 2007; Wong 2012). Results show that vowel distinction is significantly correlated with year of birth ($F(28,1) = 13.34$, $p < 0.001$), with younger speakers favoring smaller distances between vowels (figure 2).

This apparent time correlation also obtains within sample subsets of females ($F(14,1) = 4.97$, $p < 0.04$) and males ($F(12,1) = 12.56$, $p < 0.004$), with no difference between them (figure 3). However, while the correlation similarly obtains among the subset of Chinese Americans ($F(14,1) = 19.77$, $p < 0.001$), it does not obtain for the subset of European Americans ($F(12,1) = 0.07$, $p = 0.8$).

The general correlation between year of birth and Euclidean distance supports the interpretation of the token-based SPEAKER:VOWEL interaction effects as merger-in-progress. However, while the token-based models found no effects of speaker ethnicity, the mean-based model shows a significant interaction between year of birth and ethnicity. The lack of an apparent time correlation in the EA subsample does not seem to be due to the merger having gone to completion for that group, nor does it seem that all EAs are maintaining a distinction. Rather, it seems that differences among

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

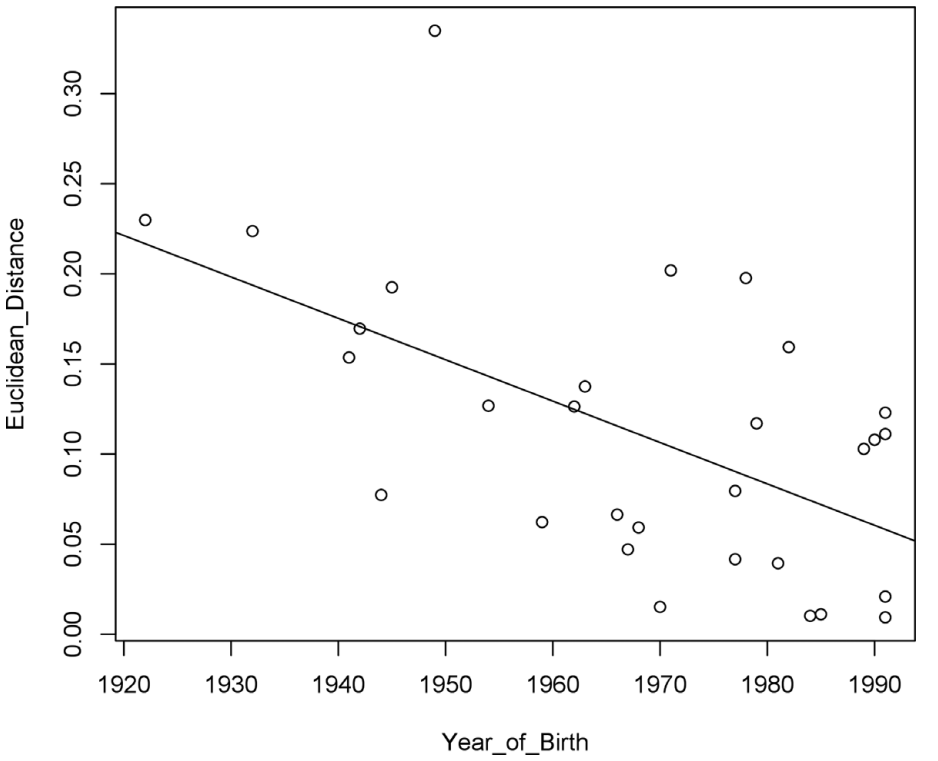


Figure 2. Speaker year of birth with respect to the Euclidean distance between COT and CAUGHT midpoints in normalized F1/F2 space

European Americans are fairly great, but are simply not correlated with age – at least not in this dataset, which admittedly does not have any speakers born before the 1940s.

To balance out the subsamples, the two oldest speakers, 'Lou' and 'Enid', were removed. The apparent time correlation was still significant for the subset of Chinese Americans ($F(13,1) = 13.87$, $p < 0.003$), and a model of merger that includes an interaction effect of between year of birth and ethnicity significantly improves on a model of only main effects ($F(27,1 = 10.64$, $p < 0.003$). This result is particularly interesting given that Moonwomon (1991) showed merger-in-progress among European Americans in the Sunset District, specifically. Something appears to have interrupted the progress of her change in apparent time; this will be discussed further in section 5.

3.3 Wordlist data

Due to a technical error, one of the 30 speakers ('JoJo') was without a wordlist. The remaining 29 speakers produced 428 tokens, 171 for COT and 257 for CAUGHT. Data were modeled in the same way as the interview data, with the additional independent

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

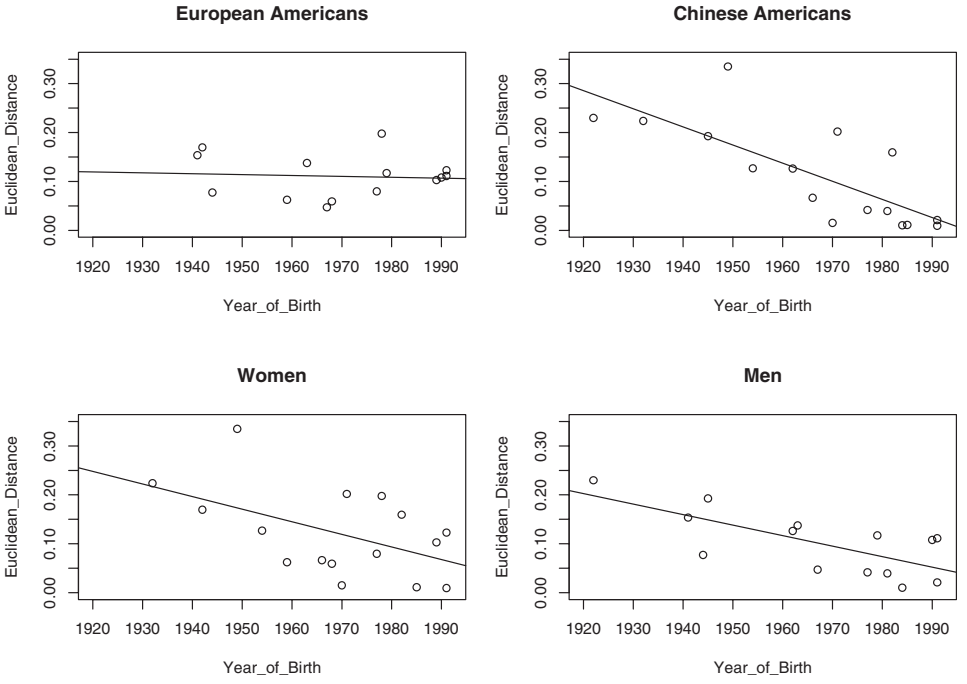


Figure 3. Speaker year of birth with respect to the Euclidean distance between a speakers' average COT and CAUGHT midpoints, in normalized F1/F2 space, by speaker subset (European Americans and Chinese Americans; women and men)

variable of vowel duration. Following environment also had fewer levels (no open syllables or following bilabials), and following fricatives only occurred for CAUGHT.

The effects predicting F1 are VOWEL and VOWEL:YOB, and the effects predicting F2 are FOLLOWING, VOWEL, YOB and VOWEL:YOB. The new factor, vowel duration, does not distinguish COT from CAUGHT, and is not a significant predictor for either F1 or F2. Overall, the wordlist results are similar the interview results. One difference is that following phonological environment is not significant predictor for F1, and, for F2, following velars do not disfavor fronted vowels ($t = 0.632$). Another difference is that year of birth is a significant *main* effect for F2 ($t = 2.671$). In other words, regardless of whether the vowel is COT or CAUGHT, younger speakers produce fronter vowels in the wordlist condition than older speakers do.

Euclidean distance measures of vowel means were also calculated for the wordlist data ($EucD_{wd}$) and compared to those of the interview data ($EucD_{int}$). One question is to what extent speakers may be exhibiting near-merger, defined as a greater vowel distinction in interview speech than wordlist speech (see Labov *et al.* 1991). The absolute value of the difference between the two conditions indicates the extent of 'style' effects for each speaker ($EucD_{diff}$).

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

The results show that 24 of the 29 speakers did not differ much between the two contexts (mean $EucD_{diff} = 0.061$). For these 24 speakers, all within two standard deviations of the mean, $EucD_{int}$ and $EucD_{wd}$ were highly correlated (Pearson's $r(22) = 0.81$, $p < 0.001$). Adding the other 5 speakers significantly weakened the correlation (Pearson's $r(27) = 0.68$, $p < 0.001$), indicating that those 5 showed a significant shift between speech conditions. Four of them showed an effect of vowel distinction, where the Euclidean distance was greater in the wordlist reading. The remaining speaker, 'Mary', showed the opposite effect, with $EucD_{wd}$ (0.049) much smaller than $EucD_{int}$ (0.198), suggesting near-merger. Ten of the 29 speakers show some increase in vowel coalescence in the wordlist context, but besides Mary only one speaker, 'Molly' ($EucD_{wd} = 0.099$, $EucD_{int} = 0.202$), shows a sizeable difference (just below two standard deviations from the median). These two near-merger style-shifters also happen to be the two speakers who produce flip-flop.

4 Introducing 'flip-flop'

4.1 Defining flip-flop

This section focuses on two speakers whose exceptional production patterns raise compelling questions for theories of vowel merger. They are outliers, defined here with respect to their random intercepts. This definition exploits the potential of random intercepts to act as indicators of individual speaker engagement in language change:

> When modeling the likelihood that a speaker will produce an innovative linguistic variant, the leaders of the change in progress would have positive intercepts and the speakers who produce more conservative variants would have negative intercepts. In terms of the value of the intercept, the closer the value is to 0, the less that speaker diverges from the overall pattern captured by the model. (Drager & Hay 2012: 62)

The mixed effect models (section 3) generated random intercept values for each speaker, for each vowel, for both F1 and F2. Speakers with positive intercepts for F1 favor lower vowels and speakers with positive intercepts for F2 favor fronter vowels. The larger the value for the random intercept, the more the speaker diverges from the general model. Here, speakers whose random intercept is more than two standard deviations from the median are considered outliers.

First, consider two statistically 'typical' speakers. 'Cheri' (figure 4) has random intercepts within one standard deviation of the median for all four measures (F1 and F2 for both COT and CAUGHT). She shows some overlap between vowel nuclei, but generally displays a pattern of distinction ($EucD_{int} = 0.170$). 'Monica' (figure 5) is another speaker with all random intercept values near the median. She shows advanced overlap between low back vowel nuclei and is most likely merged ($EucD_{int} = 0.009$).

Twenty-five more speakers are like Cheri and Monica in that their random intercept values are within two standard deviations of the median. This leaves three speakers who are outliers: 'Jenny' (CAUGHT F1, −2StDev), 'Mary' (CAUGHT F2, +2StDev) and

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

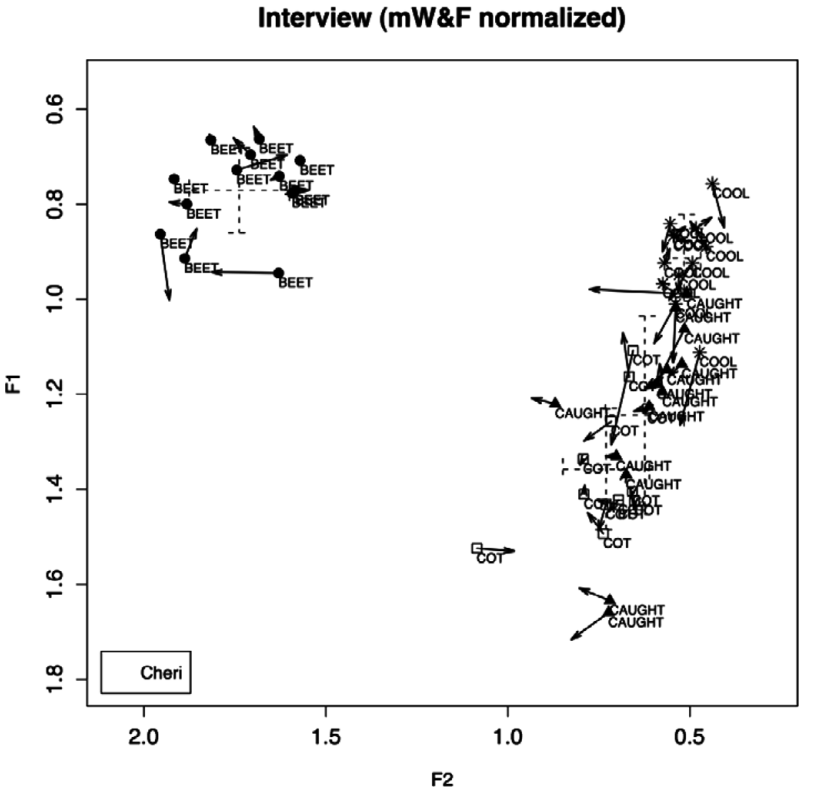


Figure 4. The BEET, COOL, COT and CAUGHT vowels as produced by 'Cheri' (female, Irish American, born in 1942), in the interview condition

'Molly' (CAUGHT F1, +2StDev and COT F1, −2StDev). Jenny has an unusually high CAUGHT vowel, Mary has an unusually fronted CAUGHT vowel, and Molly has both an unusually low CAUGHT vowel and an unusually high COT vowel. Jenny's outlier status is straightforwardly the result of her having the most conservative vowel system in the sample (figure 6), with the largest distinction between COT and CAUGHT ($EucD_{int}$ = 0.335) and with CAUGHT in its historical position. In contrast, Mary (figure 7) and Molly (figure 8) are outliers not because they are conservative but because they show flip-flop; Mary's CAUGHT is *fronter* than COT and Molly's CAUGHT is *lower* than COT.

Random intercepts, alone, do not distinguish between an outlier with a very conservative pattern (Jenny) and an outlier with flip-flop (Mary and Molly). While a positive random intercept generally corresponds to speakers with the most advanced realizations of a change (Drager & Hay 2012), for a merger-in-progress this is only true when the speaker also shows low values for vowel distinction. For example, for the F2 of CAUGHT, Monica has a random intercept of 0.003 (small, but positive) and the lowest $EucD_{int}$ value in the sample (0.009), whereas Jenny has a random intercept of −0.024 and the highest $EucD_{int}$ value in the sample (0.335). Mary and Molly are exceptional in

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

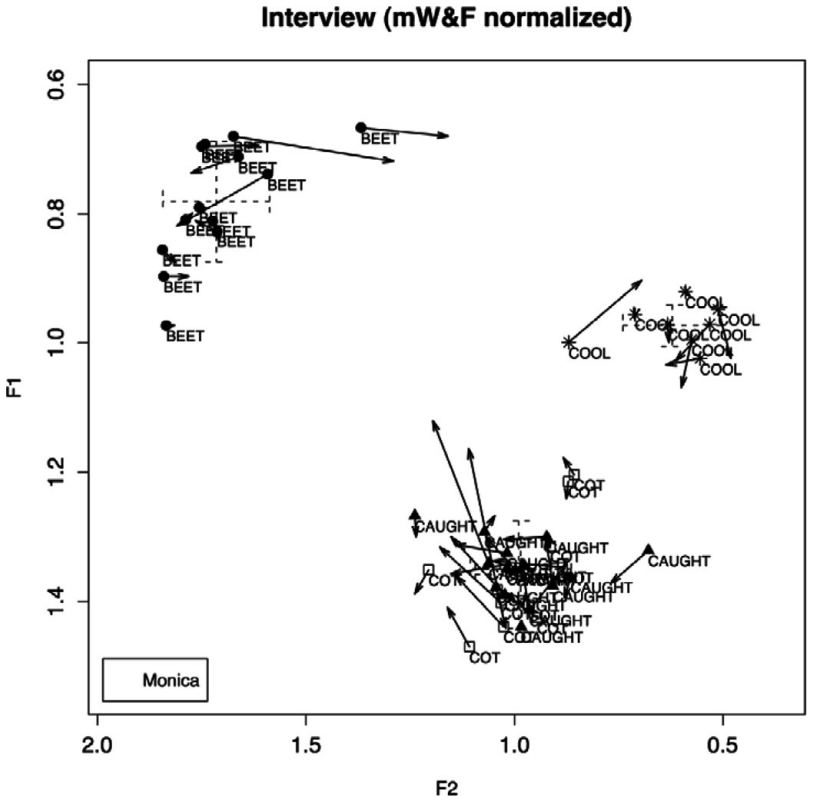


Figure 5. The BEET, COOL, COT and CAUGHT vowels as produced by 'Monica' (female, Chinese American, born in 1991), in the interview condition

that they have high positive random intercept values (Mary, CAUGHT F2 = 0.072; Molly CAUGHT F1 = 0.063) but also high $EucD_{int}$ values (0.202 and 0.198, respectively). The co-occurrence of extremely 'innovative' pronunciations with the maintenance of vowel distinction is what constitutes a flip-flop pattern in cases of merger-in-progress. Note that this definition is agnostic to the phonetic dimension(s) on which flip-flop occurs; it is not necessarily the case that all contrastive distinctions (in this case, both F1 and F2) are (completely) inverted, just that flip-flop be manifest in at least one dimension.

4.2 Flip-flop as near-merger

A speaker has a near-merger when they produce a distinction between vowel classes but perceive no distinction. In the absence of perception data, it is not possible to know for certain if Mary and Molly do perceive COT and CAUGHT to be the same vowel. However, a related feature of near-merger is when a speaker produces a greater

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

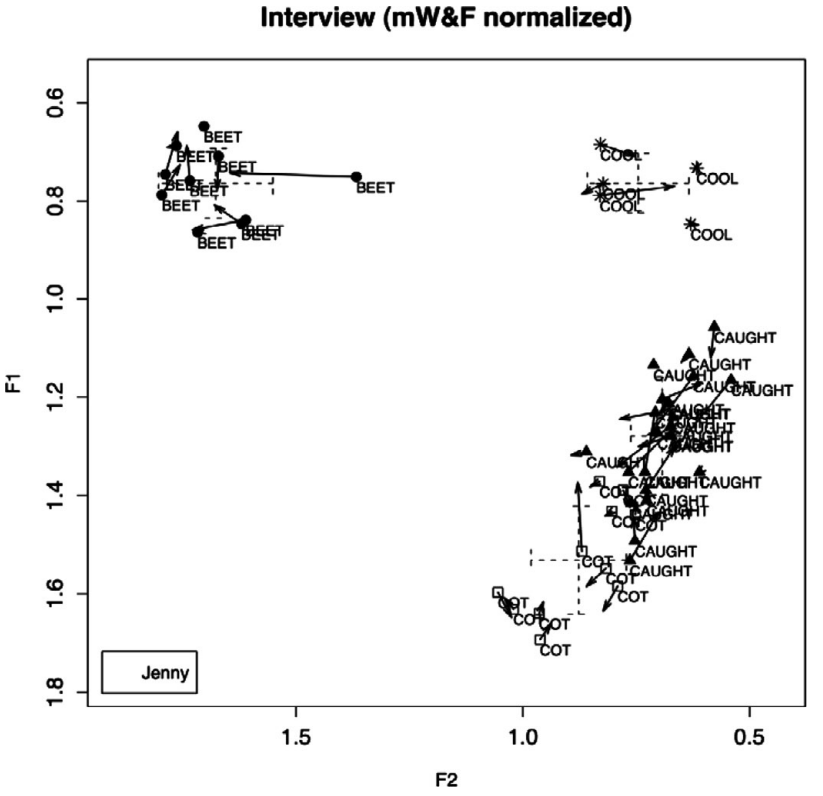


Figure 6. The BEET, COOL, COT and CAUGHT vowels as produced by 'Jenny' (female, Chinese American, born in 1949), in the interview condition

distinction in interview speech than in minimal pair wordlists (e.g. the instance of 'Dan Jones', *ibid.*). As stated in section 3, Mary and Molly show this the most clearly; both speakers (and Mary in particular) appear merged in the wordlist condition (figures 9 and 10). Paralinguistic cues and metalinguistic commentary from both speakers further suggest that neither perceives a distinction. During the wordlist task, Mary explicitly states that she cannot distinguish the vowels and has been made aware of that fact by others (1). Note that *dawn–Don* was the first COT–CAUGHT minimal pair on the list.

(1) . . . *dawn*, *Don* <pause> . . . that's one that I always get <unintelligible> on the East Coast, *Erin* and *Aaron* are the same name and *Dawn* and *Don* are the same name.

Mary then laughs immediately following the pair *cot–caught*, the next key pair after *dawn–Don*. She does not laugh after any other pair. Her laughter about *cot–caught* may be cueing her awareness in (1) that some speakers ('on the East Coast') do make a distinction, but that she does not.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

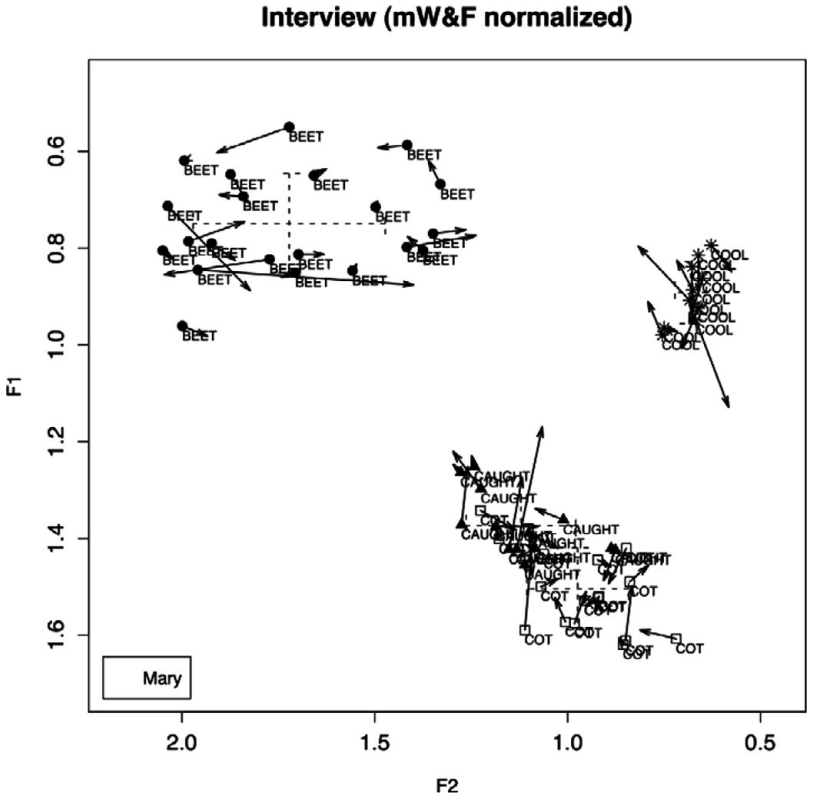


Figure 7. The BEET, COOL, COT and CAUGHT vowels as produced by 'Mary' (female, Irish American, born in 1978), in the interview condition

Molly's only metalinguistic comment during the entire wordlist task follows a key COT–CAUGHT pair (2).

(2) . . . *stalk stock* <pause> It's like a tongue-twister!

At the point of uttering (2), Molly had reached the 22nd of 80 pairs. She had been reading the list relatively quickly, thereby producing more false starts than most speakers, which may account for the content of her comment. The fact that the comment itself follows a key COT–CAUGHT pair might be a coincidence, or it might indicate a heightened awareness of difficulty for these particular pairs. The paralinguistic cue that supports this interpretation is pause length; she pauses for much longer before (*mean*, 826ms) and after (*mean*, 556ms) a minimal pair containing a CAUGHT vowel (N = 6) than in between all other minimal pairs (*mean*, 328ms, N = 32). Pause duration is also much more variable following all CAUGHT pairs (*StDev,* 615ms) than in other contexts (*StDev,* 265ms). Longer pauses are also found in or adjacent to the homonymic pairs, *two–too*

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

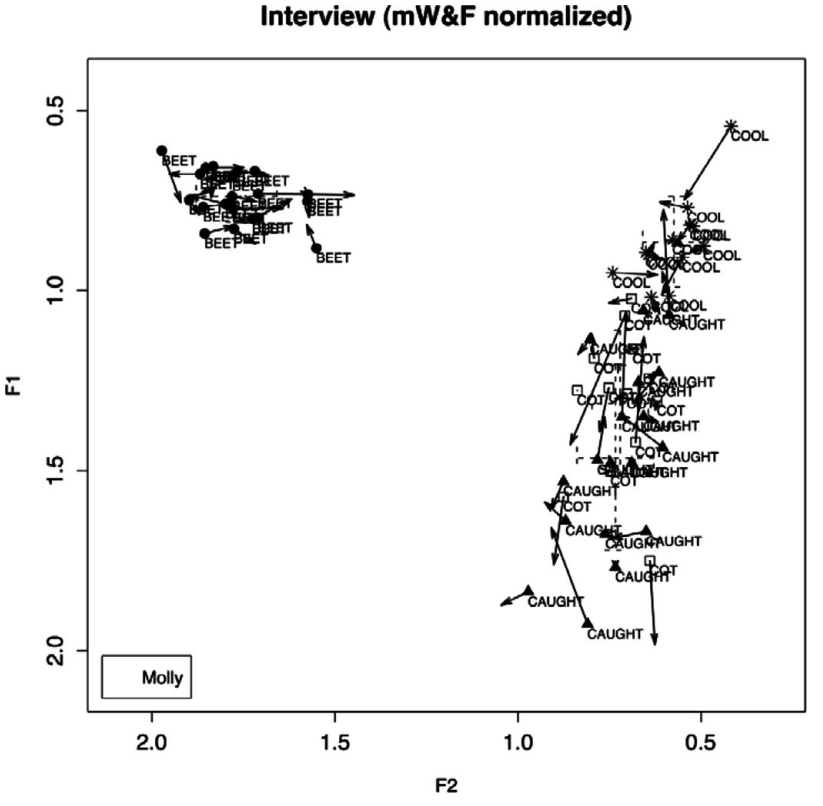


Figure 8. The BEET, COOL, COT and CAUGHT vowels as produced by 'Molly' (female, Chinese American, born in 1971), in the interview condition

and *do–dew* (*mean,* 751ms), suggesting that Mary is treating the COT–CAUGHT pairs more as homonyms than as minimal pairs.

The reduction in vowel distinction in the wordlist condition, combined with some metalinguistic and paralinguistic evidence, suggests that the two speakers who produce flip-flop are producing a distinction they do not perceive. In the absence of more direct evidence from perception, it is hypothesized that flip-flop will always coincide with near-merger, in an individual. Note that this does not help solve questions about near-mergers more generally, specifically how individuals who do not perceive a distinction nonetheless manage to produce one. Perhaps what we do see here only complicates the picture further: some near-merged speakers not only maintain a distinction, but they do so even when the phonetic cues to that distinction are in some way reversed.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

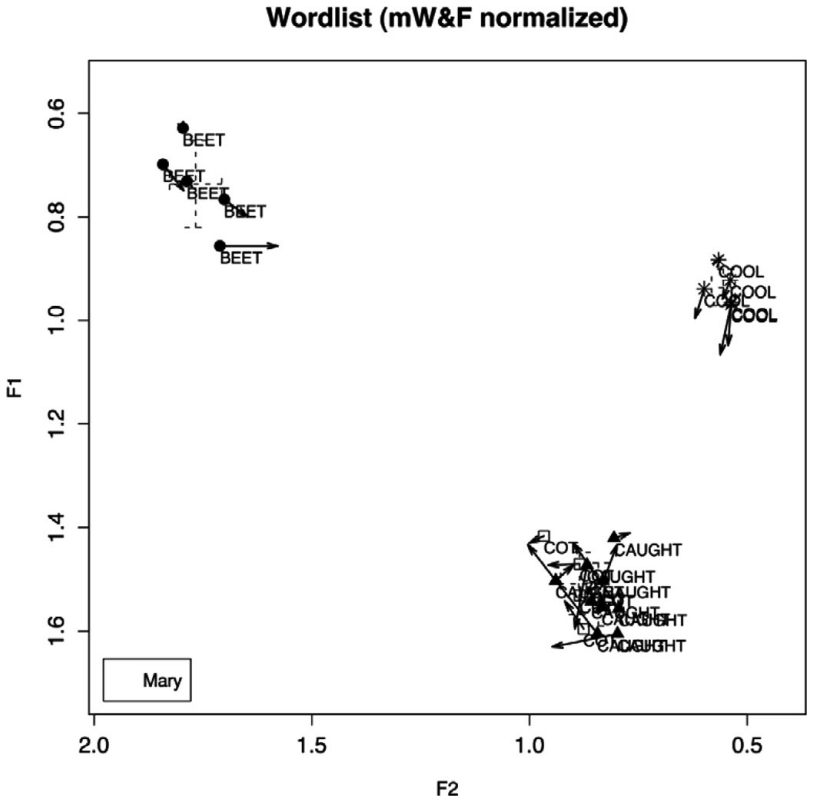


Figure 9. The BEET, COOL, COT and CAUGHT vowels as produced by 'Mary' (female, Irish American, born in 1978), in the wordlist condition

5 Analysis: accounting for 'flip-flop'

5.1 Who flip-flops? A social account

What factors make a speaker likely to produce flip-flop? At one level, this question gets to the heart of the relationship between production and perception: are these speakers misperceiving their vowel targets, and if so, why is the production consequence of this misperception only evident in free speech? Since perception data are unavailable for these speakers, the remainder of this article considers what ethnographic detail may reveal about the social qualities of the kind of person who is most likely to flip-flop. Since flip-flop is very rare, the intent here is not to predict its occurrence. Rather, when flip-flop does occur, attending to social factors of its speaker may provide useful insight into both the social structures of the community and the progress of the sound change itself.

This analysis draws on sociophonetic applications of indexicality theory (Silverstein 2003; Eckert 2008a), suggesting that part of the explanation for flip-flop lies in the social

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

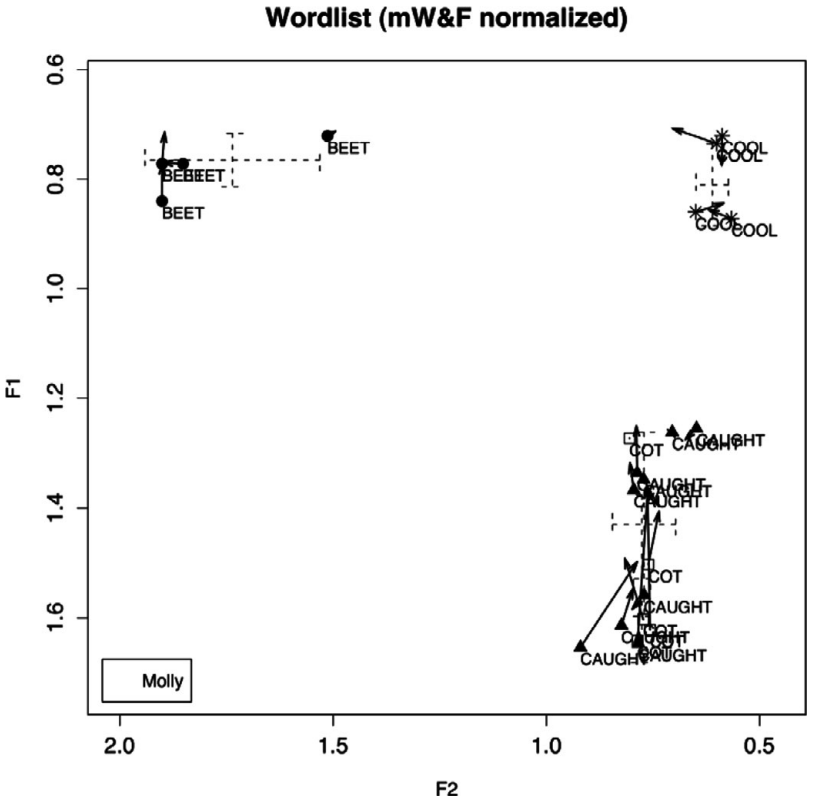


Figure 10. The BEET, COOL, COT and CAUGHT vowels as produced by 'Molly' (female, Chinese American, born in 1971), in the wordlist condition

meaning(s) indexed by the extreme phonetic quality produced by the flip-flopped variant (here, the CAUGHT vowel). As argued elsewhere (Wong & Hall-Lew under review) the variable quality of CAUGHT is available as a resource of authentication (Bucholtz 2003). In San Francisco's Sunset District, its value differs with respect to speaker age and ethnicity, by virtue of the shift in signifiers of local authenticity (section 1). Higher and backer forms of CAUGHT may index an older authenticity while lower and fronter productions may index newer local meanings. Flip-flop may be one consequence of a speaker's negotiation of indexes of local authenticity that were changing rapidly during her youth.

5.2 Mary

Mary is an unusual interviewee for a number of reasons (see also Hall-Lew forthcoming). She is very metalinguistically aware (being one of the only speakers to explicitly comment on the COT–CAUGHT merger; (1)) and very socially aware. In

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

discussing her time at middle school,[11] she spontaneously offered a thorough analysis of communities of practice, listing the names and habits of each group. She justified her level of detail by describing her youth as a series of attempts to cross various social boundaries. As an adult, she prides herself on having an extremely diverse friendship network and the ability to adapt easily to different social groups, a skill she characterizes as 'San Franciscan'.

Mary's multifaceted identity is apparent in her biographic profile, in how she talks about herself, and in how other people talk about her. Mary is Irish American; on her father's side she is fifth-generation Irish San Franciscan and a third-generation Sunset District resident. For a neighborhood that did not exist before the twentieth century, this is a remarkable level of historicity. In (3), Mary reflects on the older Irish Catholic authenticity of the Sunset, which she associates with her father's generation:

(3) . . . like a remnant Sunset really. I don't- you know it's like, um, more like my dad's era of Sunset than mine, where there are ton of like cops' kids and firemens' kids and carpenters' kids and all that stuff, and they go to Catholic school all the way through . . .

Mary frames her own experience of the Sunset ('mine') in contrast to her father's ('my dad's era'), and draws on reference to occupation type and schooling as points of contrast. Both are, ultimately, references to ethnic change: civil service jobs and Catholic schools are associated with Irish identity for nearly all residents older than Mary. The contrast between these two 'eras', Mary's and her father's, is a reflection of the real demographic shifts of the neighborhood (section 1). The shift from Irish dominance to Asian dominance increased most rapidly in the 1980s, when Mary was growing up. The majority of Mary's peers are not Irish Catholics, and their parents' professions are highly diverse.

At the same time, being Irish is a fundamental aspect of Mary's identity. The community clearly dominated her earliest years of socialization; as an adult, she recognizes that fact, but not without ambivalence:

(4) [<Interviewer:> How do you identify with your Irish side? Do you feel like that's a part of how you present yourself and identify?]
Yeah. Yeah, being a native, for sure, yeah, um. <pause> . . . My soccer team . . . our sponsor bar's in the Sunset . . . a total like dirt bag, <laughter> like Irish immigrant bar, you know? At any given time at least fifty percent of the people are first generation, you know, huge brogues . . . there'll be like a crock pot in back making, like, a pot roast in the middle of the afternoon <laughter> . . . People come in between a wedding and a . . . reception, and, like, hang in this, like, dark bar in the middle of the day. I'm like, 'you're so weird' . . . But that's also, you know, um, there are people that I know from growing up who that's like more their day-to-day than not, you know?

Mary first interprets my question about ethnicity as a question about 'being a native'. The mere recognition of this link is rare among San Franciscans her age. However,

[11] Grade 6–8, approximately ages 12–14.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

immediately following explicit identification as Irish, she constructs a distance between herself and certain Irish cultural practices ('a crock pot in back making, like, a pot roast in the middle of the afternoon'). The pejorative 'dirt bag' and the reenacted quote, 'I'm like, 'you're so weird', convey her disassociated stance. Yet her critique ends with a return to alignment with Irish identity through her recognition that these same people are familiar from her childhood. This discursive moment is merely one instance of Mary's negotiated and variable orientation to an ethnic identity; the bulk of the interview is comprised of similar examples. Her repeated emphasis on the fluidity of her identity practices, in particular her frequency of style-shifting, may be connected to her production of low back vowel flip-flop.

One of the earliest social motivations for style-shifting for any individual is transitioning from home to school. Neighborhood schools exemplified the ethnic shift happening in the wider Sunset District community. Even in 2008, the largest private schools were Catholic, and the Irish population initially dominated their instruction and attendance. In contrast, the largest local public[12] schools acquired multiethnic student populations earlier. In the 1980s, some already had minority white (and in some cases, majority Chinese) student populations, in part because of the neighborhood demographic change but also because of a city bussing system that brought in students from other areas of San Francisco. Among the 30 speakers analyzed here (excluding Lou, who was mainly schooled in China), across all ages, 73 percent of the Chinese Americans but only 46 percent of the European Americans attended public schools. Another 25 percent of Chinese Americans and 38 percent of the European Americans attended some mixture of both public and Catholic schooling. The three remaining speakers – two older Irish Americans and Molly – fit Mary's description of her father's 'remnant Sunset', having only attended Catholic schools (2). Mary, unlike the Sunset Irish before her, had a mixed schooling experience, switching from public to private between middle school and high school. She struggled with social acceptance much more at the public school than in the private school. Mary's mother (who is not native to San Francisco, but a long-time resident) discusses this in (5).

(5) [My daughters] had pretty multiethnic friendships. [Mary] particularly had, um, friends that were not Caucasian. When she was at that [public] school her group of choice was Black girls, but that didn't last very long. They wouldn't, um, take her in, if that makes any sense. . . . Her Black friends got in trouble because they were hanging out with her. And the same thing happened, Asian. [Mary] would be disturbed to lose certain friends because, because it became too painful for them to cross that line.

As a young adolescent, Mary clearly struggled to gain acceptance among peer groups whose ethnicities differed from her own, in a context where those differences were so strong they often prohibited the possibility of friendship. One might imagine that this traumatic experience was part of what motivated Mary to then switch to an Irish Catholic high school, where she could abandon the goal of constructing an identity

[12] Where 'public' refers to state-funded, non-denominational education and 'private' refers to schooling that charges tuition fees (either religious or non-denominational).

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

that crossed ethnic lines. However, she states that the main motivation to attend that school was out of family tradition (following her father and grandfather). Her attempt to cross racial boundaries also continued to persevere into adulthood. In talking about her daughters in 2008, Mary's mother says:

(6) Both girls are very proud that they are from here, and that ... they have an active interpersonal life with people from everywhere ... [Mary] is very <with emphasis> involved with the Asian culture here.

Mary's self-descriptions corroborate her mother's, both in terms of her childhood history and adulthood. Toward the end of her interview, when discussing the relationship between language and San Francisco identity (prompted by the interviewer's questions), Mary says:

(7) You get to pull like, anything you want out of the different influence[s] you had growing up, and it comes off natural ... I was about as cartoonishly, you know, upwardly-mobile-white-girl as you get, but I could slip into how I talked when I was in middle school and people were like 'Whoa, how do you know how to talk like that?' you know and I was like, 'Oh <laughter> all I wanted to be when I was thirteen was a five-foot-tall Asian girl who could breakdance!' <laughter>

In (7), Mary claims facility over remarkable linguistic style-shifting abilities and attributes them to the attempted acquisition of 'Asian' styles around the age of 13. While she asserts a certain level of recognized fluency (on the part of her contemporary interlocutors) she frames her 13-year-old bids for identity as aspirations, not direct reflections of her friendship network. One imagines this 13-year-old Irish American girl experimenting with variable linguistic styles as part of a general attempt to gain entry into social spaces dominated by non-Irish ethnic groups. At first glance it may be perplexing that a fourth-generation Sunset District Irish American, an individual with unquestionable claims to authenticity, would struggle with social acceptance in a local school. But her experience reflects the dramatic shift in the neighborhood's economy of symbolic value, away from Irish Catholic social practices and toward an 'Asian'-dominated semiotic space.

Without ratified access to the relevant communities of practice, then, the occurrence of flip-flopped forms on Mary's part is, perhaps, not surprising. The quantitative interspeaker results suggested that merger-in-progress has been active in the mainstream, Chinese American community, and not as active in the increasingly minoritized European American community. Since Chinese Americans constitute a clear majority among the Sunset District 'Asian' population, and since the merger is being achieved through movement of the CAUGHT vowel, it may be that more advanced CAUGHT tokens are a component of 'Asian' styles, or just new local persona more generally. Perhaps an advanced CAUGHT vowel was just one small part of the stylistic package that indexed 'five-foot-tall Asian girl[s] who could breakdance'. Or, perhaps conservative productions of CAUGHT indexed Irish Catholic or just traditional local persona more generally – social identities that Mary wished to avoid.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

In this scenario, flip-flop is posited to be one result of a lifetime of negotiated (and often rebuffed) bids for authenticity: a speaker with the biographical historicity (Coupland 2003) to claim one authenticity, but who aims instead for a newer system of local value, perhaps 'overshooting' her target. While her childhood context provided her distinct COT and CAUGHT phonemes, as she grew older she may have settled on a production pattern that reflects a constant (re-)negotiation of self. However, despite her mother's negative characterization, as an adult Mary views her lifelong bid for multiple identities as generally successful. In (8) she discusses her only time away from San Francisco, which was four years at Cornell University.

(8) [The] East Coast always kind of like appealed to me and I think actually part of why I transitioned so well from like scene to scene as an adult is because I did have this kind of like disparate upbringing in that regard.

'[I]n that regard' refers to an immediately prior description of her mother's family who live on the East Coast and who Mary describes as initially intimidating and strange because they were 'more WASP-y than I even knew'. While she frames White-Anglo-Saxon-Protestant culture as something more foreign than the Asian cultures at her middle school, she also draws on this as further evidence of her cross-cultural background and resultant style-shifting skill, a skill she sees as beneficial, allowing her to 'transition so well' between different social groups.

5.3 Molly

Molly also experienced an unusually high degree of social negotiation in childhood and adolescence, emerging as an adult who is explicitly proud of her ability to style shift. In contrast to Mary, Molly experienced life in the Sunset District as a Chinese American. Like many of the Chinese Americans in the sample, she was Chinese-dominant (in Mandarin) until age 5. As a child born in 1971 she would have been a member of a clear minority (whereas as an adult she is now part of the majority). This experience was amplified by the fact that all her schooling was at Irish Catholic institutions. When asked about what the Sunset was most known for during her childhood, Molly says: 'Irish. <laughter> The Irish.' Molly is the youngest Chinese American in the sample to recognize a relationship between Irish identity and local authenticity. Her interview also recounts memories of anti-Asian violence, while younger Chinese Americans, who grew up in an Asian majority, were often only vaguely aware of that violence (see Sharma 2011 for parallel generational differences in London). When asked about childhood friends, Molly begins by saying that 'most' of her friends were from school, but then ends by emphasizing the friendships she fostered outside school, which she characterizes in terms of ethnicity:

(9) When I was growing up, most of my friends, they were from my school . . . I had a lot of friends that were more- <pause> like, friends that I met through my family's, like, my family's, you know, daughter or son. And I had friends that lived in the neighborhood, like, a couple of friends from across the street . . . School was mainly

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

> Irish dominated or, um, [it] was like a mixed group. But you know my friends outside of school were mainly Asian, just because, um, the nature of the way the community was here, kind of like, Asians kinda stuck together.

Like Mary, aspects of Molly's interview indicate that having multiple friendship groups may have been a strategic response to struggles to gain social acceptance. In (10), she mentions 'getting picked on a lot' in contrast to her younger sister who attended the same schools but is 'very Irish' (i.e. culturally Irish, since she is also of Chinese heritage).

(10) She married an Irish guy; she's very Irish . . . Her experience in the Sunset is very different from my experience, and most of her friends, I- no, I would say maybe fifty percent of her friends, are Asians, but a lot of her best friends growing up in the Sunset were Irish . . . We went to the same schools. It just so happened- I think there was more discrimination growing up when- when I was growing up, than three years later, when she was . . . I remember getting picked on a lot.

Molly clearly did not have her sister's ease adapting to the culture of their schools. Although there is only a three-year age difference between them, Molly attributes their different experiences to rapid social change, and in other parts of the interview continues to frame social relations in the Sunset District during her childhood as being more ethnically segregated and with 'more discrimination' than in following years. Flip-flop may be one result of Molly employing various strategies to cope with the lack of social acceptance. This analysis predicts that Molly's sister would not to show flip-flop. In other words, even though both girls acquired their English phonology at school, I would predict that the mere fact of having been a Chinese American in an Irish Catholic school in the 1970s is an insufficient predictor of flip-flop production. Rather, the similarities between Molly and Mary suggest that the motivation is more to do with the struggle to gain acceptance and the strategic response of style-shifting (11).

(11) I adapt to wherever I go. So, when I was with my public school friends I had this kind of Chinglish talk, like, 'hey man' <laughter> you know, you know 'what's up' you know? <laughter> And it was, like, totally different from the way I talked to my private school friends, which was very proper.

In (11), Molly reinforces the association between 'public school' and Chinese ethnicity (where 'Chinglish' refers to Chinese-influenced English) and introduces the association between 'private school' and talking 'proper'. Neither linguistic style is presented as better than the other, but Molly presents herself as someone with equal command over both. Like Mary, Molly presents style-shifting as a skill, an ability to 'adapt' which she has carried into her adult life. Molly lived away from San Francisco between the ages of 22 and 30, moving between various international cities before moving back to the Sunset after 9/11. In (12), she frames her experience again in terms of style-shifting and the command over a broad linguistic repertoire:

(12) [W]hen I moved to Hong Kong, um, I st- I started speaking British English . . . After being there for like three years my sister came to visit me and she's like 'why do you

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

> talk like an FOB?'[13] and I said I <laughter> I can go British-British accent, like, with a Chinglish accent. . . . I had to be able to speak to the locals so I had to change my English . . . And then I wanted to move to LA, I <laughter> started speaking LA English, so kind of more like a Valley Girl kind of accent just, just because everyone else there did.

One of the parallels between Molly (12) and Mary (7) is that both report other people's surprise or confusion at their extent of style-shifting. That they show remarkable linguistic flexibility as well as an unusual vowel pattern may not be a coincidence.

6 Discussion

COT and CAUGHT are clearly undergoing merger-in-progress in San Francisco's Sunset District. While COT remains low and central, CAUGHT is fronter and lower among younger speakers, with a decrease in the acoustic distance between the vowels. In the Sunset District neighborhood, Chinese Americans appear to be leading the change. The manifestation of community linguistic norms (merger-in-progress) among the group who now defines community social norms (Chinese Americans) is perhaps unsurprising, although a notable finding for sociophonetic studies of Asian Americans and sound change (Ito 2010; Wong 2012). Furthermore, given the ethnic focus of previous studies, the present findings indicate that the San Franciscan resistance to merger (Labov *et al*. 2006) may be more characteristic of European Americans than the city as a whole. Crucially, the interaction between age and ethnicity for merger suggests an indexical relationship between vowel and local meanings of ethnicity (see Wong & Hall-Lew under review). As sound change was moving through the Sunset District, so was a dramatic shift in local identity from an 'Irish parish' to a 'New Chinatown'. If the rates of sound change and social change were connected, then more distinct vowels might index older community norms and the merged vowel might index newer community norms. Since COT is stable in apparent time, CAUGHT would be the vowel indexing these social meanings. One account of the 'interruption' of Moonwomon's (1991) merger-in-progress among European Americans would thus be the emergence of local value for conservative vowel variants.

Flip-flop occurs from speakers' realization of CAUGHT being further along one dimension of a phonetic trajectory of a sound change than the merged speakers' realizations along that same dimension. This results in the flip-flopped vowel's formant values (for example) being higher than those of even the youngest speakers.[14] Flip-flop suggests a speaker's orientation toward the social meaning(s) associated with advanced CAUGHT. It would result either from a speaker's attempt to avoid a back/high CAUGHT, or to produce a front/low CAUGHT, or both. One area for future work would be to

[13] FOB stands for 'Fresh Off the Boat', originally a pejorative term for immigrants (but see Hall-Lew forthcoming).

[14] Note that the notion of *hypercorrection* is avoided here because there is no single notion of 'correct' pronunciation in this particular context. However, the notion may be historically relevant here (see DeCamp 1972).

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

examine stance taking during utterances with the most extreme productions. It is also noteworthy that flip-flop occurred here in speech directed to an interlocutor with the low back vowel merger;[15] future research would benefit from comparing these data to interview and wordlist data conducted by a speaker with the traditional low back vowel distinction, to test for accommodation effects.

A number of factors further complicate these findings. Recall that the interaction between age and ethnicity did not obtain for the token-based analysis of CAUGHT for either F1 or F2, alone, but only in combination, across averages. Furthermore, the discussion of 'advanced CAUGHT' is complicated by the fact that Mary flip-flops with respect to anteriority, whereas Molly flip-flops with respect to height. Labov (1994: 359) notes that, for near-mergers, the maintained phonetic difference 'is most often an F2 difference, instead of a combination of F1 and F2'. One might expect flip-flop to similarly only occur along F2, especially given DeCamp's (1953: 555) description of CAUGHT variability in San Francisco as consisting of 'fronted allophones'. However, other work (Wong & Hall-Lew under review) would predict flip-flop along the F1 dimension, the argument being that CAUGHT lowering is more salient than fronting; small frequency differences in F1 are more perceptible than the same differences in F2. Relatedly, Di Paolo's (1992) 'Speaker 251' flip-flopped with respect to F1 but not F2. Complicating matters, the merger Mary exhibits in wordlists entails not only a decrease in F2 (versus her interview speech) but also an increase in F1 (to accomplish merger). She is flip-flopping with respect to anteriority but also varying with respect to height. Lastly, Mary's realization of flip-flop only concerns CAUGHT, while Molly's pattern is due to both vowels. Since COT is not undergoing change in the community, Molly's pattern is unusual for (perhaps) independent reasons.

Together, these differences indicate that Mary and Molly have rather distinct patterns of flip-flop. With only two speakers, it is not possible to account for these differences. Perhaps they are aiming to index different social meanings, or perhaps they are aiming for the same meaning but only one of them accurately perceives/produces the relevant phonetic cues. The major difference between them is ethnicity, and therefore early linguistic exposure. Perhaps differences in realization of flip-flop are due to their variable access to the relevant norms during language acquisition. Indeed, it is interesting to consider why flip-flop is *not* obviously more likely for a member of the 'outgoing' ethnic group than it is for the 'incoming' ethnic group. Unfortunately, this only raises more questions than can be answered here. The discussion instead attends to the broader fact that both speakers are clear outliers, regardless of the details of their outlier status.

Macro-social similarities suggest possible social motivations for flip-flop. First, both speakers came of age during a major demographic shift in their community, and in the middle of a sound change-in-progress. They are a generation younger than the speakers with distinction, who think of the community as 'Irish', and a generation older than

[15] Thanks to Jenny Cheshire and Kevin Stadler for different aspects of this observation.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

the speakers with merger, who think of the community as 'Asian'. But it may also be notable that both are women. The other Western US English speaker exhibiting COT–CAUGHT flip-flop (Di Paolo's 'Speaker 251') was also female, and would be about the same age as the speakers here (born around 1970). It is not unwarranted to make a more general prediction that women might be more likely than men to exhibit flip-flop, especially given examples of change-in-progress that show women producing more advanced variable realizations (e.g. Labov 2001).

Despite coming from 'opposite' ethnic communities, both Mary and Molly show ambivalence toward their early experiences with ethnicity. Mary has spent much of her life bidding for affiliations with non-Irish social groups while maintaining a flexible and ambivalent association with her inherited Irish identity. Molly also experienced ambivalence toward Irish identity, but for exactly the opposite reasons. Coupland (2003: 428) asks, '[i]f individuals cannot securely inherit authenticity from the social circumstances of their birth and socialisation, how can they achieve it?' Both speakers appear to have developed strategies of authentication to navigate experiences in early adolescence. Frequent style-shifting and repertoire expansion allowed them to navigate social boundaries. Here, Mary is seen shifting between the interview and wordlist contexts more than any other speaker. A more thorough approach to style might very well reveal that both speakers have highly variable linguistic repertoires (Sharma 2011). In short, flip-flop may be but one linguistic consequence of a lifetime of active negotiation between conflicting local authenticities.

7 Conclusion

A community undergoing merger-in-progress typically shows individuals producing vowels at greater or lesser degrees off overlap, with retained distinctions in the direction of the historically prior vowel distribution. Flip-flop effects are, therefore, as surprising as they are rare. However, when they do appear, they may be both socially and linguistically revealing. As a co-occurring phenomenon with near-merger, the age of the flip-flop speaker may indicate the fulcrum point in the trajectory of the shift. But unlike near-merger, knowing the life experiences of the flip-flop speaker may point to the social meanings indexed by that sound change. In short, although flip-flop in near-merger rarely occurs, attending to the life histories and personal narratives of its speakers provides an opportunity to see how individuals navigate the interaction between social change and sound change. Ethnographic insight, therefore, can provide the necessary detail to build more comprehensive and predictive models of the range of phenomena associated with vowel mergers.

Author's address:
Linguistics & English Language
University of Edinburgh
Dugald Stewart Building
3 Charles St

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Edinburgh EH8 9AD
UK
Lauren.Hall-Lew@ed.ac.uk

References

Baayen, R. H. 2010. *languageR: Data sets and functions with 'Analyzing Linguistic Data: A practical introduction to statistics'.* R package version 1.0. http://CRAN.R-project.org/package=languageR.

Bloch, Matthew, Shan Carter & Alan McLean. 2011. Mapping the 2010 US Census. *The New York Times.* http://projects.nytimes.com/census/2010/map. Accessed 13 December 2011.

Boberg, Charles. 1999. The attitudinal component of variant in American English foreign (a) nativization. *Journal of Language and Social Psychology* 18, 49–61.

Boberg, Charles. 2009. The emergence of a new phoneme: Foreign (a) in Canadian English. *Language Variation and Change* 12, 355–80.

Boberg, Charles. 2010. *The English language in Canada: Status, history and comparative analysis.* Cambridge: Cambridge University Press.

Boersma, Paul & David Weenink. 2012. *Praat: Doing phonetics by computer*, version 5.3.04. Retrieved 12 January 2012 from www.praat.org.

Bucholtz, Mary. 2003. Sociolinguistic nostalgia and the authentication of identity. *Journal of Sociolinguistics* 7(3), 398–416.

Coupland, Nikolas. 2003. Sociolinguistic authenticities. *Journal of Sociolinguistics* 7(3), 417–31.

DeCamp, David. 1953. The pronunciation of English in San Francisco. PhD dissertation, University of California, Berkeley.

DeCamp, David. 1959. The pronunciation of English in San Francisco (part 2). *Orbis* 54–77.

DeCamp, David. 1972. Hypercorrection and rule generalization. *Language in Society* 1(1), 87–90.

Di Paolo, Marianna. 1992. Hypercorrection in response to the apparent merger of (oh) and (a) in Utah English. *Language and Communication* 12, 267–92.

Di Paolo, Marianna, Malcah Yaeger-Dror & Alicia Beckford Wassink. 2011. Analyzing vowels. In Marianna Di Paolo & Malcah Yaeger-Dror (eds.), *Sociophonetics: A student's guide*, 87–106. London and New York: Routledge.

Drager, Katie & Jennifer Hay. 2012. Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change* 24, 59–78.

Eberhardt, Maeve. 2008. The low-back merger in the Steel City: African American English in Pittsburgh. *American Speech* 83, 284–311.

Eckert, Penelope. 2003. Sociolinguistics and authenticity: An elephant in the room. *Journal of Sociolinguistics*, 7(3), 392–7.

Eckert, Penelope 2008a. Variation and the indexical field. *Journal of Sociolinguistics* 12, 453–76.

Eckert, Penelope. 2008b. Where do ethnolects stop? *International Journal of Bilingualism* 12, 25–42.

Fabricius, Anne H., Dominic Watt & Daniel Ezra Johnson. 2009. A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change* 21, 413–35.

Fridland, Valerie. 2004. The spread of the *cot/caught* merger: A look at region and race. Paper presented at NWAV33, Ann Arbor, MI.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Godfrey, Brian J. 1988. *Neighborhoods in transition: The making of San Francisco's ethnic and nonconformist communities.* Berkeley: University of California Press.
Gordon, Matthew. 2006. Tracking the low back merger in Missouri. In Thomas E. Murray & Beth Lee Simon (eds.), *Language variation and change in the American Midland: A new look at 'Heartland' English*, 57–68. Philadelphia: John Benjamins.
Hall-Lew, Lauren. 2009. Ethnicity and phonetic variation in a San Francisco neighborhood. PhD dissertation, Stanford University.
Hall-Lew, Lauren. 2010. Improved representation of variance in measures of vowel merger. *Proceedings of Meetings on Acoustics (POMA)*, 9(1).
Hall-Lew, Lauren. 2011. The completion of a sound change in California English. *Proceedings of ICPhS XVII.* [CD-ROM]
Hall-Lew, Lauren. Forthcoming. Chinese social practice and San Francisco authenticity. In Véronique Lacoste, Jakob Leimgruber & Thiemo Breyer (eds.), *Indexing authenticity: Sociolinguistic perspectives*. Berlin and New York: de Gruyter.
Hall-Lew, Lauren. Under review. 'I went to school back East . . . in Berkeley': San Francisco English and changes in authenticity. In H. Samy Alim, Patricia Baquedano-Lopez, Mary Bucholtz & Dolores Ines Casillas (eds.), *Vox California: Cultural meanings of linguistic diversity.*
Herold, Ruth. 1997. Solving the actuation problem: Merger and immigration in eastern Pennsylvania. *Language Variation and Change* 9, 165–89.
Hinton, Leanne, Sue Bremner, Hazel Corcoran, Jean Learner, Herb Luthin, Birch Moonwomon & Mary Van Clay. 1987. It's not just Valley Girls: A study of California English. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13, 117–27.
Ito, Rika. 2010. Accommodation to the local majority norm by Hmong Americans in the Twin Cities, Minnesota. *American Speech* 85(2), 141–62.
Irons, Terry Lynn. 2007. On the status of low back vowels in Kentucky English: More evidence of merger. *Language Variation and Change* 19, 137–80.
Irvine, Judith & Susan Gal. 2000. Language ideology and linguistic differentiation. In P. V. Kroskrity (ed.), *Regimes of language: Ideologies, polities, and identities*. Santa Fe: School of American Research Press, 35–84.
Johnson, Daniel Ezra. 2007. Stability and change along a dialect boundary: The low vowels of Southeastern New England. Dissertation, University of Pennsylvania.
Kurath, Hans & Raven I. McDavid. 1961. *The pronunciation of English in the Atlantic states [PEAS].* Ann Arbor: University of Michigan Press.
Labov, William. 1994. *Principles of linguistic change*, vol. 1: *Internal factors*. Oxford: Blackwell.
Labov, William. 1998. The three dialects of English. In M. D. Linn (ed.), *Handbook of dialects and language variation*, 39–81. San Diego: Academic Press.
Labov, William. 2001. *Principles of linguistic change*, vol. 2: *Social factors*. Oxford: Blackwell.
Labov, William. 2007. Transmission and diffusion. *Language* 83, 344–87.
Labov, William. 2010. *Principles of linguistic change*, vol. 3: *Cognitive and cultural factors*. Oxford: Blackwell.
Labov, William, Mark Karen & Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3, 33–74.
Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress: Report on National Science Foundation Contract GS-3287*. Philadelphia: US Regional Survey.
Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology, and sound change*. New York and Berlin: Mouton de Gruyter.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Laguerre, Michel S. 2005. The globalization of a panethnopolis: Richmond District as the New Chinatown in San Francisco. *Geojournal* 64, 41–9.
Majors, Tivoli. 2005. Low back vowel merger in Missouri speech: Acoustic description and explanation. *American Speech* 80, 165–79.
McCall, Leslie. 2005. The complexity of intersectionality. *Signs* 30, 1771–1800.
Moonwomon, Birch. 1991. Sound change in San Francisco English. PhD dissertation, University of California, Berkeley.
Moore, Emma & Robert Podesva. 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38, 447–85.
Nycz, Jennifer. 2005. The dynamics of near-merger in accommodation. In Sylvia Blaho, Luis Vicente & Erik Schoorlemmer (eds.), *Proceedings of ConSOLE XIII*, 273–86.
Nycz, Jennifer. 2011. Second dialect acquisition: Implications for theories of phonological representation. PhD dissertation, New York University.
Pederson, Lee. 2001. Dialects. In John Algeo (ed.), *The Cambridge history of the English language*, vol. 6: *English in North America*, 253–90. Cambridge: Cambridge University Press.
Plichta, Bartlomiej. 2006. *Akustyk: A free Praat plug-in for sociolinguists.* http://bartus.org/akustyk/.
R Development Core Team (2010). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria, www.R-project.org.
Rauniomaa, Mirka. 2003. Stance accretion. Paper presented at the Language, Interaction, and Social Organization Research Focus Group, University of California, Santa Barbara, February.
Sharma, Devyani. 2011. Style repertoire and social change in British Asian English. *Journal of Sociolinguistics* 15(4), 464–92.
Terrell, Tracy. 1975. The merger of the low vowel phonemes in the English of Southern California: A preliminary report. Paper presented at the American Dialect Society.
Thomas, Erik R. 2011. *Sociophonetics: An introduction*. Basingstoke and New York: Palgrave Macmillan.
Thomas, Erik R. & Tyler Kendall, 2007. NORM: The Vowel Normalization and Plotting Suite: An online tool for sociophonetic vowel normalization. http://NCSLAAP.lib.NCSU.edu/Tools/NORM./ Accessed: 4 December 2012.
Ungaretti, Lorri. 2012. *Stories in the sand: San Francisco's Sunset District, 1847–1964*. San Francisco: Balangero Books.
Wells, J. C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
Wong, Amy Wing-mei. 2012. The lowering of raised-THOUGHT and the low-back distinction in New York City: Evidence from Chinese Americans. *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NWAV 40* 18(2), 157–66.
Wong, Amy & Lauren Hall-Lew. Under review. Regional variability and ethnic identity: Chinese Americans in San Francisco and New York City. In Malcah Yaeger-Dror & Lauren Hall-Lew (eds.), *New perspectives on the concept of ethnolect.*
Yaeger-Dror, Malcah, Lauren Hall-Lew & Sharon Deckert. 2003. Situational variation in intonational strategies. In P. Leistyna & C. Meyer (eds.), *Corpus analysis: Language structure and language use*, 209–24. Amsterdam: Rodopi.
Yu, Alan. 2007. Understanding near mergers: The case of morphological tone in Cantonese. *Phonology* 24, 187–214.

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press

Appendix

Best models for F1 and F2 for the interview speech data:

F1	Estimate	Standard error	t
Intercept	1.1928	1.6735	0.713
VOWEL = CAUGHT	−3.9518	1.3564	−2.913
FOLLOWING = bilabial	0.0245	0.0172	1.424
FOLLOWING = fricative	0.0088	0.0186	0.469
FOLLOWING = nasal	−0.0801	0.0186	−4.303
FOLLOWING = open	−0.0307	0.03	−1.025
FOLLOWING = velar	0.0023	0.0142	0.164
YOB (linear)	0.0001	0.0009	0.164
VowelCAUGHT:YOB (linear)	0.002	0.0007	2.877
F2	**Estimate**	**Standard error**	**t**
Intercept	0.5380	1.6377	0.329
VOWEL = CAUGHT	−5.4142	1.1161	−4.851
FOLLOWING = bilabial	−0.0205	0.0171	−1.200
FOLLOWING = fricative	−0.0439	0.0181	−2.421
FOLLOWING = nasal	−0.0636	0.0172	−3.691
FOLLOWING = open	−0.0365	0.0340	−1.075
FOLLOWING = velar	−0.0376	0.0150	−2.498
YOB (linear)	0.0002	0.0008	0.205
VowelCAUGHT:YOB (linear)	0.0027	0.0006	4.813

Best models for F1 and F2 for the wordlist speech data:

F1	Estimate	Standard error	t
Intercept	2.6565	1.7558	1.513
VOWEL = CAUGHT	−7.8520	1.3674	−5.742
YOB (linear)	−0.0006	0.0001	−0.643
VowelCAUGHT:YOB (linear)	0.0039	0.0001	5.671
F2	**Estimate**	**Standard error**	**t**
Intercept	−0.1789	1.1857	−0.151
VOWEL = CAUGHT	−3.6014	0.8595	−4.190
FOLLOWING = fricative	−0.0377	0.0134	−2.806
FOLLOWING = nasal	−0.0271	0.0108	−2.515
FOLLOWING = velar	0.0078	0.0123	0.632
YOB (linear)	0.0005	0.0006	0.838
VowelCAUGHT:YOB (linear)	0.0018	0.0004	4.137

https://doi.org/10.1017/S1360674313000063 Published online by Cambridge University Press