WORKING PAPERS

IN

LINGUISTICS

The notes and articles in this series are progress reports on work being carried on by students and faculty in the Department. Because these papers are not finished products, readers are asked not to cite from them without noting their preliminary nature. The authors welcome any comments and suggestions that readers might offer.

Volume 44(3)
April
2013

DEPARTMENT OF LINGUISTICS
UNIVERSITY OF HAWAIʻI AT MĀNOA
HONOLULU 96822

DEPARTMENT OF LINGUISTICS FACULTY
2013

Victoria B. Anderson
Andrea Berez
Derek Bickerton (Emeritus)
Robert A. Blust
Lyle Campbell
Kenneth W. Cook (Adjunct)
Kamil Deen
Patricia J. Donegan (Co-Graduate Chair)
Katie K. Drager
Emanuel J. Drechsel (Adjunct)
Michael L. Forman (Emeritus)
John H. Haig (Adjunct)
Roderick A. Jacobs (Emeritus)
William O’Grady
Yuko Otsuka
Ann Marie Peters (Emeritus, Co-Graduate Chair)
Kenneth L. Rehg (Chair)
Lawrence A. Reid (Emeritus)
Amy J. Schafer
Albert J. Schütz, (Emeritus, Editor)
Jacob Terrell

THE EFFECT OF A SINGLE FORMANT ON DIALECT IDENTIFICATION[1]

JAMES GRAMA

Labov (2001:167–68) makes the claim that English speech communities use F_2 in vowels to establish social identity, while they use F_1 chiefly for the cognitive differentiation of vowel phonemes. However, little work has been done to address whether this observation holds in perception. By using a forced-choice, matched-guise experiment, this paper investigates whether variations in a single formant can shift perceptions of a speaker's regional origin. Results suggest that when the F_1 of DRESS is low, the vowel is more reliably rated as Californian, suggesting that depending on the vowel, both formants may be important in the perception of social identity.

1. INTRODUCTION. Labov (2001:167–68) argues for the social pre-eminence of the second formant and states that English speech communities seem to use "F_1 for cognitive or categorical differentiation, and differences in F_2 for establishing social identity." While vowel height has been shown in the phonetics literature to be perceptually important in the categorization of vowel phonemes (Ladefoged and Broadbent 1956), little work in sociophonetics has explicitly investigated whether specific formants (e.g., F_2) play a role in conveying social identity. Vowel stimuli in sociolinguistic experiments are usually manipulated in both F_1 and F_2 (Niedzielski 1999; Hay and Drager 2010). This makes it difficult to determine the effect that a single formant has on dialect identification. Using electronically manipulated vowel stimuli of the short front vowels KIT, DRESS, and TRAP[2] from speakers of California English (CalE), this study investigates Labov's claim about the social pre-eminence of the second formant in perception[3] and experimentally tests whether listeners identify dialect differences using F_2 more reliably than F_1.[4]

Perception studies in phonetics and sociolinguistics have attempted to describe how vowels are categorized and how they relate to social meaning. While the categorization of vowels is based largely on the phonemic inventory of a language (Walley and Flege 1999), vowel perception is also influenced by other top-down information, such as the expected dialect of a speaker (Rakerd and Plichta 2010). Sociolinguistic perception studies have provided valuable insight into the social meaning of linguistic variables and can inform production studies by experimentally testing what people report they believe about speech patterns. These types of studies have investigated listener perceptions of dialect differences, often with respect to vowels, and the extent to which listeners can identify a speaker's regional origin (Purnell et al. 1999; Clopper and Pisoni 2004; Baker, Eddington, and Nay 2009), as well as how the perception of vowels can be affected by social information (Niedzielski 1999; Hay et al. 2006; Hay and Drager 2010). Research investigating a speaker's ability to identify regional origin has shown that more widespread exposure to varieties leads to a better ability to perceive differences among dialects. Clopper and Pisoni (2004), for example, found that listeners who had lived in several different states ("army brats") were more accurate in identifying the speaker's dialect area than listeners who had lived only in

[1] I wish to thank Katie Drager and my other qualifying paper committee members, Kamil Deen and Amy Schafer, for their support and insight. I would also like to thank Christian Koops for his advice and suggestions. All errors of fact and interpretation are my own.

[2] This paper uses the lexical sets established by Wells (1982).

[3] Research in sociolinguistics has demonstrated that individuals have some knowledge of sociolinguistic variables and use this knowledge during the perception of speech (Ladefoged and Broadbent 1957; Fridland and Okamoto 2009).

[4] Though this paper explicitly makes reference to formants, there is an argument to be made that phonological features might more generally be indexed to a social category (cf. Scharinger and Lahiri 2010). While this paper acknowledges that a featural interpretation is possible, it makes explicit reference to formants as the acoustic variable in question.

Indiana (“homebodies”). Furthermore, Baker, Eddington, and Nay (2009) demonstrated that residents of Utah were more accurate in identifying a Utah dialect than non-Utahns, and that locals focused more consistently on a specific phonetic parameter than non-Utahns. However, knowing additional social information about a speaker can also heavily influence a listener’s perception. Niedzielski (1999) found that listener responses to stimuli from the same dialect varied based on the dialect the listeners expected to hear, and Rakerd and Plichta (2010) demonstrated that listeners alter their phonetic category boundaries depending on the dialect of a preceding carrier phrase. Additionally, Hay et al. (2006) showed that listeners’ ability to differentiate distinct tokens of a merger in progress in New Zealand English (NEAR and SQUARE) varied according to the age, gender, and perceived socioeconomic status of a person presented in a photograph.

This study uses results from a forced-choice, matched-guise experiment to show that lowered and retracted realizations of DRESS and TRAP are perceptually recognized as markers of a California way of speaking. Furthermore, shifting F_1 in DRESS is sufficient to alter perceptions of a speaker’s regional origin. This highlights the importance of both formants (F_1 and F_2) in perception and suggests crucially that either F_1 or F_2 may be more correlated with social factors, depending on the vowel.

2. Literature Review

2.1 Formants in perception. Labov (2001:167) noted for his data that the correlation between social factors and vowels is almost entirely concentrated in F_2. He claimed that in general, “nothing is to be gained by introducing F_1 into the analysis of sociolinguistic constraints” (167). This argument is largely based on his observations about near mergers of vowel classes and the amount of variance explained by F_1 in his Language Change and Variation in Philadelphia (LCV) project. First, he (1994) demonstrated that near mergers (i.e., vowels that are merged in perception but not yet in production) show significant overlap in F_1 in production, and when a distinction is made between two nearly merged vowels, it is in F_2. Furthermore, his (2001) statistical models in the LCV project show that F_2 is the best predictor of variation in his data set. In fact, very little additional variance is explained when both F_1 and F_2 are combined in his regression model (165–67), with the notable exception of the PRICE lexical class when followed by a voiceless consonant. Based largely on these two points, he concluded that F_1 is used chiefly for “cognitive or categorical differentiation, and differences in F_2 [are used] for establishing social identity,” though he admitted that further inquiry should be made to investigate the way these observations correlate with “pyscho-acoustic mechanisms” (168).

Labov’s claim is corroborated in the phonetics literature, at least regarding the importance of F_1 for categorizing vowel phonemes. While perception experiments in phonetics have documented that F_1 and F_2 are sufficient together for categorical identification of vowels (Fry et al. 1962:172), varying F_1 has been shown to strongly affect the categorical perception of vowel quality in English (Ladefoged and Broadbent 1957). Furthermore, Fry et al. (1962) demonstrate that the perceptual boundaries among vowels are less sharply defined than the boundaries among consonants. Despite this, varying both F_1 and F_2 along a continuum results in non-continuous perception of vowels that is heavily affected by top-down information (Fry et al. 1962). For example, the perception of a vowel can be affected by whether the vowel is in a word that has lexical status (Walley and Flege 1999), the vowel’s proximity to other vowels (Ladefoged and Broadbent 1957), and the perceived identity of the speaker (Johnson 1990).

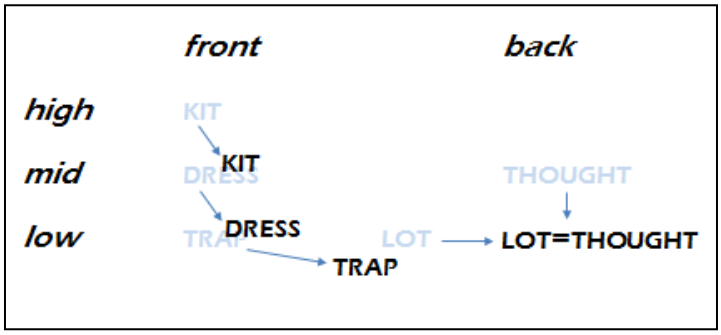
Socio-perceptual studies have been effective in identifying how the knowledge of social variables affects the perception of vowels (e.g., Hay et al. 2006), but the experimental stimuli used in these studies are often varied along both F_1 and F_2. This work manipulates formants to create a continuum between two vowels in which each token varies in both F_1 and F_2. Because both formants are manipulated, experiments that use these types of experimental stimuli are not equipped to investigate whether one formant has a greater effect than another formant on listener perception.

The present study manipulates formants within vowels, resulting in tokens with various combinations of frontness and backness. By creating experimental items that acoustically vary in only one formant and by asking listeners to identify the dialect region of the voice, this study investigates whether one formant

(specifically F_2, as Labov predicts) is more reliably correlated with social information than the other. The following section describes the relevant phonetic variables of the dialect used to test this question.

2.2 SHORT FRONT VOWELS IN CALIFORNIA ENGLISH (CALE). California was chosen as the dialect group used to implement this experiment. First, CalE is currently undergoing a change in progress involving the short front vowels KIT, DRESS, and TRAP (Hagiwara 2005; Eckert 2008; Kennedy and Grama 2012), where these vowels are realized as lower (i.e., exhibit an increased F_1 value) and backer (i.e., exhibit a decreased F_2 value) in pre-obstruent position relative to other American English varieties (see figure 1, adapted from Grama and Kennedy 2009). The short front vowels in CalE are particularly conducive to addressing which of the first two formants is more reliably correlated with social information in perception because both formants are realized differently from their canonical American English values. These vowels are therefore better suited to investigate this question than a change that involves mainly only one formant (e.g., a change in progress involving the fronting of GOOSE, which is largely reflected in F_2). Therefore, the listener is perhaps equally likely to attend to either F_1 or F_2 when making judgments about a speaker's regional origin. There is also reason to believe that the short front shift in California is a change that is below the level of consciousness (i.e., a change from below). Californians do not seem to comment explicitly on it as a feature of California English (Kennedy and Grama 2012), but lowered short front vowels do appear in some parodies of social stereotypes associated with California (e.g., *valley girls* and *surfer dudes*) (Hinton et al. 1987; Kennedy and Grama 2012). Furthermore, this lowering and retraction parallels the Canadian shift (Clarke, Elms, and Youssef 1995), itself a change from below (Labov 2002).

FIGURE 1: Graphic representation of the short front chain shift in CalE (adopted from Grama and Kennedy 2009)



2.3 THE MATCHED-GUISE METHOD IN DIALECT IDENTIFICATION. A methodology that has been successful in isolating the contextual meaning of specific linguistic variables is the matched-guise technique (Purnell et al. 1999; Levon 2007; *inter alia*). In this methodology, a single speaker is presented to the listener in multiple guises or contexts (e.g., first saying the word *fishing* with an alveolar nasal [n] and then saying it with a velar nasal [ŋ]), and the listener reports his/her intuitions about the speaker, such as the speaker's origin, sexuality, ideals, appearance, values, or personality. While this methodology has traditionally used a multi-dialectal or multi-lingual speaker to produce the different guises (e.g., Purnell et al. 1999), speech manipulation software allows even more methodological control and the ability to investigate single variables at a time (Levon 2007; Campbell-Kibler 2008; Kirtley 2011). By manipulating a single phonetic variable within an utterance, a researcher can create two (or more) different versions of the utterance that differ in the target variable. This effectively controls for all other linguistic variables in the utterance. Because the guises are "matched" for speaker and acoustic context, any difference in reported intuitions about a speaker can confidently be attributed to the phonetic variable in question.

Important for this study, matched-guise experiments have been successful in showing that social judgments can be made even when there is very little phonetic information present (van Dommelen and Moxness 1995; Purnell et al. 1999), such as when the listener is only exposed to only a single-word (e.g., the word *hello* in Purnell et al. 1999). This study capitalizes on this observation and uses single-word

tokens as stimuli to create a large amount of experimental control across the test stimuli. What follows is a discussion of the methodology used to create the test stimuli and the way the experiment was implemented.

3. METHODOLOGY

3.1 STIMULI CREATION. Stimuli were created to test two questions: (1) are low, retracted realizations of the short front vowels markers of CalE? And (2) do listeners use F_2 more reliably than F_1 when evaluating social factors? To address these questions, stimuli were created using data taken from read wordlists produced by ten speakers (5 male, 5 female) collected during fieldwork sessions in various locations in California. To test whether listeners evaluate lowered, backed tokens of short front vowels spoken by non-CalE speakers as Californian, similar wordlist data were taken from two speakers born in Hawai‘i.[5] This yielded a total of 12 speakers (6 female, 6 male) who were used as critical stimuli during the experiment.[6] A breakdown of demographic information about the speakers can be found in table 1. Wordlist data were chosen to effectively and efficiently control for phonological context, duration, and prosodic variables, and to ensure that the experiment would isolate as much as possible the effect of varying a single formant. Single-word tokens were also chosen as the stimuli for this experiment so that a larger number of speaker voices might be incorporated. All experimental items took the form hVd (e.g., *hid*, *head*, *had*).

TABLE 1: Overview of speakers used to create stimuli.

Speaker region	# of Males	# of Females	Mean age
California	5	5	23.7
Hawai‘i	1	1	19.5
Mean age	22.5	23.5	23
Total # of speakers	6	6	12

To determine the degree to which the short front vowels needed to be manipulated, the F_1 and F_2 values were measured at the midpoint of vowels KIT, DRESS, TRAP, FLEECE, FACE, and LOT from monosyllabic tokens in pre-obstruent or word-final positions for all CalE voices. No tokens were analyzed before nasals, as pre-nasal contexts influence the color of short front vowels in CalE (see Eckert 2008). All tokens were analyzed using Akustyk (Plitchta 2010). All vowels were measured during their steady state at roughly the midpoint of the vowel's duration. FACE was the only exception to this, as it is diphthongal for California speakers; it was therefore measured at the midpoint of the steady state of its nucleus. In total, 83 short front vowels from Californians were analyzed. These values were then averaged for each speaker over the total number of instances of each vowel. Using Akustyk, each speaker's vowel space was plotted using the extracted formant values. This plot was then compared against the relative values of each speaker's vowels, and how shifted the vowels sounded to the author (himself a native CalE speaker). Based on this information, it was determined that all 10 speakers from CA exhibited the short front shift.

Before the formant values were manipulated, all tokens were converted from stereo to mono, normalized for intensity at 50 dB (SPL), and resampled to 11,025 Hz. The tokens were then bordered with silence: 250 ms before the onset of each token and 50 ms of silence following each token.

[5] These speakers were extracted from the SOLIS corpus at the University of Hawai‘i at Mānoa (Drager et al. 2012).

[6] This number reflects the total number of experimental items following a naturalness rating task by eight trained phoneticians. Initially, 15 speakers were manipulated, but this naturalness rating task removed three of them (two females and one male).

As the short front chain shift in California is characterized by relatively lower and backer articulations of the short front vowels (i.e., an increase in F_1 and a decrease in F_2), all the tokens for eight of the CalE speakers were manipulated using Akustyk to yield two tokens each: one where F_1 was decreased and F_2 was left unchanged (corresponding to a higher but still back vowel), and one where F_1 was left unchanged and F_2 was increased (corresponding to fronter but still low vowel). The remaining short front vowels from two California speakers (1 female, 1 male) were manipulated so that each short front vowel was both raised and fronted. These two speakers serve as a control, as their vowels are acoustically "less Californian" in both dimensions (i.e., a high-front position relative to what is expected for younger Anglo CalE speakers).

For all native CalE speakers, manipulated vowels were constructed by using the midpoint between the F_1 of the target vowel and the F_1 vowel immediately above it in the chain series, as shown in equations (1–3). In other words, the change in F_1 for TRAP was calculated as the midpoint in Hertz between the F_1 of TRAP and DRESS, and the change in F_1 for DRESS was calculated as the midpoint between the F_1 of DRESS and KIT. As there is no reference vowel that is higher than KIT in the short front chain series, the increase in F_1 for KIT was determined as the midpoint between the F_1 for KIT and the F_1 for FLEECE.

$$(1)\ \Delta F_1(TRAP') = \frac{F_1(TRAP) - F_1(DRESS)}{2}$$

$$(2)\ \Delta F_1(DRESS') = \frac{F_1(DRESS) - F_1(KIT)}{2}$$

$$(3)\ \Delta F_1(KIT') = \frac{F_1(KIT) - F_1(FLEECE)}{2}$$

Similarly, changes in F_2 were constructed by using the midpoint between the F_2 of the target vowel and the F_2 of the vowel immediately above it in the chain series, as shown in equations (4–6). In other words, the change in F_2 for TRAP was calculated as the midpoint in Hertz between the F_2 of DRESS and the F_2 of TRAP, and the change in F_2 for DRESS was calculated as the midpoint in Hertz between the F_2 of KIT and the F_2 of DRESS. Again as there is no reference vowel higher than KIT in the short front chain series, the change in F_2 for KIT was determined by calculating the midpoint between the F_2 of FACE and the F_2 of KIT. The class FACE was chosen as an anchor relative to KIT instead of FLEECE due to a tendency for the F_2 of FLEECE to be dramatically greater than for any of the surrounding vowels. The F_2 of FACE more accurately paralleled the differences achieved by the calculations used to shift DRESS and TRAP.

$$(4)\ \Delta F_2(TRAP') = \frac{F_2(DRESS) - F_2(TRAP)}{2}$$

$$(5)\ \Delta F_2(DRESS') = \frac{F_2(KIT) - F_2(DRESS)}{2}$$

$$(6)\ \Delta F_2(KIT') = \frac{F_2(FACE) - F_2(KIT)}{2}$$

Hawai‘i speakers were used to test whether listeners evaluate lowered, backed tokens of short front vowels spoken by non-CalE speakers as Californian. For the Hawai‘i speakers, short front vowels were manipulated so that F_1 was increased by 100 Hz and F_2 was decreased by 200 Hz. If vowel position is the main phonetic parameter being used to evaluate the Californianness of the stimuli, the shifted Hawai‘i-born speakers should be identified as Californian. If the Hawai‘i speakers are not identified as Californian, it would suggest that listeners are attenuating to phonetic information other than formants when making this decision.

Each experimental item belonged to one of four possible categories, reflecting the location of each of the vowels relative to their actual manifestation (see figure 2):

(1) *highback*: F_1 was decreased (the vowel was raised) and F_2 was left unchanged.
(2) *lowfront*: F_2 was increased (the vowel was fronted) and F_1 was left unchanged.
(3) *highfront*: F_1 was decreased and F_2 was increased (the vowel was both raised and fronted).

(4) *lowback*: F_1 was increased and F_2 was decreased (the vowel was both lowered and backed).

FIGURE 2: Manipulation directions of CalE speakers (left) and Hawai‘i-born non-CalE speakers (right)

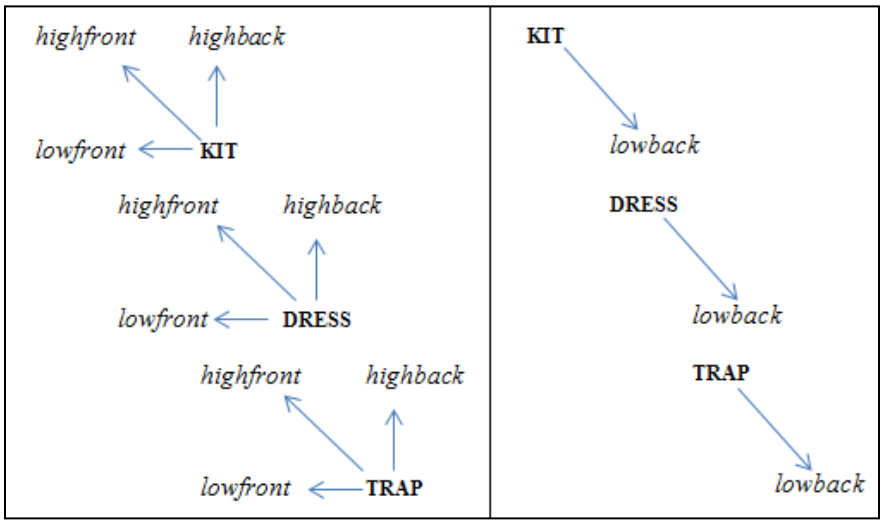
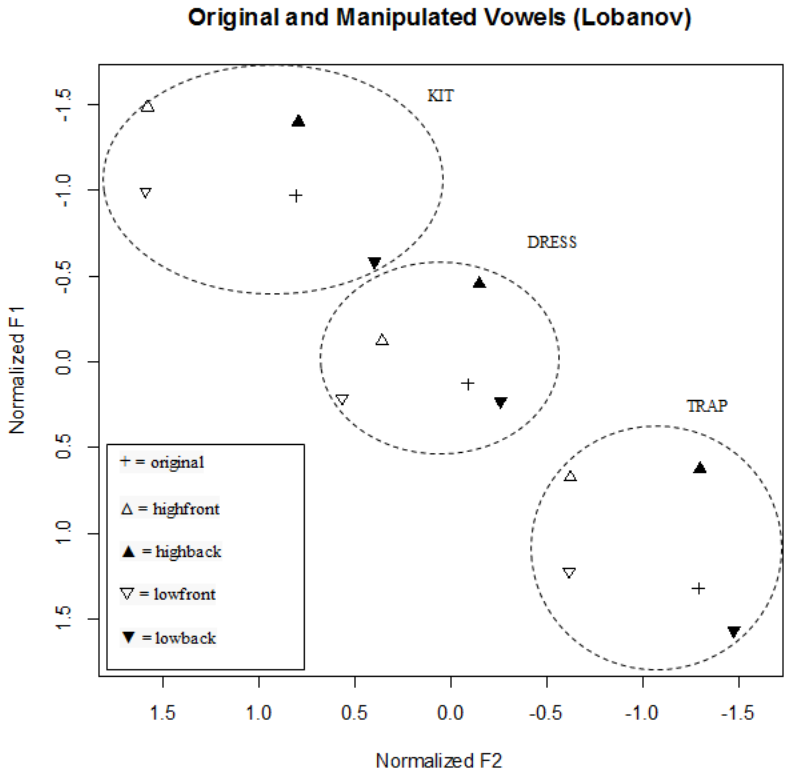


FIGURE 3: Labanov normalized mean values of original vowel token and manipulations for vowels KIT, DRESS, and TRAP in each experimental context (*highfront*, *highback*, *lowfront*, *lowback*).



A total of 60 critical stimuli (30 female, 30 male) were used for the experiment. Table 2 provides a breakdown of experimental items. Although all stimuli took the form hVd (e.g., *hid*, *head*, *had*), the trends that surface are assumed to extend to each word's lexical class more generally. The tokens in the *highback* and *lowfront* categories (*n*=48) are the main stimuli of interest, as they are aimed at teasing apart which aspect of the chain shift (i.e., lowering or backing) CalE speakers perceive to be more

Californian. If participants consistently evaluate *highback* vowels as Californian, this would be in line with Labov's (2001) observation that F_2 is more closely tied to social information. Conversely, if *lowfront* vowels are consistently evaluated as Californian, this would run counter to Labov's claim (i.e., suggesting that F_1 is used more reliably in the perception of social distinctions). The *highfront* category (n=6) aims to determine whether Californianness can be identified in the absence of the formant cues that phonetically characterize the dialect. If *highfront* tokens are accurately identified as Californian in single, isolated words, this would suggest that listeners were drawing inferences about speakers' geographical backgrounds based on some other phonetic cue (e.g., pitch, voice quality, realization of /d/, etc.). However, if listeners identify these speakers less reliably as CalE, it would suggest that both formants are important to the identification of a token as CalE. This same hypothesis is tested in a different way in the cases where vowels from Hawai'i-born, non-CalE speakers are manipulated to be part of the *lowback* category (n=6). If the members of the *lowback* category are consistently rated as non-Californians, this would suggest that the presence or absence of additional phonetic cues signaled to the listener that the voice was not CalE. If the *lowback* tokens are evaluated as Californian, this would suggest that listeners are relying on F_1 and F_2 alone when making their decisions.

TABLE 2: Breakdown of experimental items for males (M), females (F) and speakers from different regions

Token	***highback***		***lowfront***		***highfront***		***lowback***		**Total**
Speaker Region / Sex	M	F	M	F	M	F	M	F	
CalE	12	12	12	12	3	3			54
HI							3	3	6
Total	24		24		6		6		60

A total of 60 fillers (30 female, 30 male) using the same hVd form were extracted from similar wordlist data. Vowels in these fillers were evenly distributed among FLEECE, MOUTH, PRICE, and GOOSE, and one additional short-front vowel was used from each speaker. Half of these speakers were from Hawai'i (extracted from the SOLIS corpus), and the other half were personal acquaintances of the author who were from various states around the U.S. (Illinois, Arkansas, Georgia, Texas, and Idaho).

3.2 EXPERIMENTAL PROCEDURE. Participants for the perception experiment were recruited using word of mouth and the Linguistics Beyond the Classroom program, a program at the University of Hawai'i at Mānoa that gives students class credit in exchange for taking part in research. Participants were between the ages of 18 and 56 (median age = 21). All were natives of California and were judged by the author to be participating in the short front shift. Furthermore, 12 participants reported being born in Northern California and 19 in Southern California.[7] Eighteen participants reported having at least minimal exposure to a second language.

Participants were told that they would hear a number of voices from all over the United States, and that their job was to determine whether they thought the speaker was from California or not. Participants were also told to pay attention to the way the words were pronounced as opposed to the meanings of the lexical items themselves. Prior to the experimental task, participants answered a series of short demographic questions asking for information about their residence history, age, time spent living in Hawai'i, and proficiency in languages other than English. The experiment was run in a sound attenuated booth on a PC using the program *E-prime*; tokens were played through a pair of Sony® MDR-V250

[7] For the sake of this experiment, region is treated as binary. Though most participants were from larger metropolitan areas in Northern (e.g., the Bay Area) or Southern California (e.g., the greater Los Angeles area), two were from the Central Coast near San Louis Obispo. As these areas for many residents of CA are more closely associated with Southern CA, for the purposes of this experiment, these two participants were coded as residents of Southern California.

headphones. The participants were instructed to select whether they believed the speaker was Californian or not by pushing one of two buttons marked on a keyboard. Each participant took part in a practice session before the experiment. None of the stimuli that appeared in the experiment appeared in the practice session. Tokens in the practice trial were normalized along the same parameters as all other stimuli. There was no appreciable difference between the practice session and the experiment itself. No feedback information was given to the participants at any point during the experiment. Following the experiment, participants were debriefed as to the goal of the experiment in lay terms. Participants were asked immediately following the experiment what they noticed about the speech, and whether they were aware they had been listening to manipulated speech. Although some participants reported hearing the same voices more than once, no participants reported that they were aware that those voices had been electronically manipulated or changed in any way.

The experiment employed a within-subject methodology using the matched-guise technique. All tokens were randomized, but no two critical tokens for the same vowel in the same voice were heard consecutively. Participants heard single word utterances in a series of 120 stimuli. This intra-speaker presentation style allowed each token to be compared against itself.

A total of 31 participants (11 male, 20 female) heard all 120 stimuli (60 critical, 60 filler). Participants also saw each word written on the screen in front of them as they heard it. They were allotted a maximum of 10 seconds to make their decision about the speaker. Failing this, the experiment would move forward. No responses were removed from the total dataset for this or any other reason. A choice of either “Californian” or “Not Californian” was acceptable for each token; these responses were coded so that the “correct” answer was in line with the actual origin of the speaker irrespective of the manipulation. Tokens were coded for vowel type (KIT, DRESS, TRAP), speaker and speaker sex; all responses were coded for listener age, sex, place of origin within California, and whether they reported an L2.

4. RESULTS. Overall performance for correctly identifying speaker origin (Californian vs. non-Californian) of all experimental items was only slightly above chance (=60.3%), a result not altogether unsurprising given that stimuli were manipulated to sound less Californian, and that listeners have a difficult time correctly evaluating speaker origin even with non-manipulated stimuli (see overall results in Clopper and Pisoni 2004:39). However, listeners evaluated manipulated voices (i.e., all stimuli created from Californian speech as well as the two manipulated Hawai‘i speakers) as Californian more often than the fillers (manipulated: 67.5%, fillers: 51.8%). This suggests that listeners may have been making their decisions based on formants as well as other phonetic or prosodic aspects of each word.

To investigate whether stimulus type or listener’s demographic information affected responses, a logistic regression model with mixed effects (i.e., a mixed effects model)[8] was fit to listener’s binary responses (Californian or non-Californian) of all critical tokens (i.e., all manipulated tokens) with subject and item included as random effects and listener sex, age, region of origin, stimuli vowel type, and movement direction as fixed effects. Table 3 displays the results from the initial regression model. This model’s output represents how likely it is that a listener selected “Californian” based on each of the independent variables. The estimated intercept is the log odds that a listener’s response to the stimuli TRAP in the *lowfront* guise is “yes.” The estimate coefficient’s sign and how close it is to zero indicates the direction and strength of the effect, respectively. No effect surfaced for listener sex, age, or region.

[8] All models generated with R (R Core Development Team 2004).

TABLE 3: Logistic regression model (reporting treatment contrasts) of listener responses over all manipulated experimental items

All Critical Tokens	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	1.9184	0.3329	5.762	< 0.0001
Vowel.DRESS	-0.3208	0.1325	-2.420	0.0155
Vowel.KIT	-0.9098	0.1290	-7.050	< 0.0001
Movement.*highback*	-0.2102	0.1185	-1.774	0.0761
Movement.*highfront*	-0.5458	0.1795	-3.041	0.0024
Movement.*lowback*	-1.3017	0.1783	-7.300	< 0.0001

Table 3 shows that vowel type (regardless of manipulation) was important in interpreting the overall results; TRAP was more likely evaluated as Californian than either DRESS ($p = 0.01$) or KIT ($p < 0.0001$). This effect is corroborated by looking at overall percentages, where out of 496 tokens per vowel, TRAP was evaluated as Californian 392 times (79.03%), DRESS 362 times (72.98%), and KIT only 284 times (57.26%). Both of these measures suggest a picture where TRAP was most likely to be attributed to a CalE speaker, followed by DRESS and KIT. Including this response bias in the model helps to statistically control for vowel type so that other factors in the model can be considered above and beyond vowel identity.

Overall manipulation direction also surfaced as important. Recall that California voices were those voices in the *highfront*, *highback*, and *lowfront* guises, with *lowback* voices being from Hawai‘i. Both the *highfront* ($p = 0.002$) and *lowback* ($p < 0.0001$) categories were less likely to be evaluated as Californian than vowels in the *lowfront* category. Vowels in the *highback* category also trended towards being evaluated as less Californian, though the model did not return significant values for the difference between *highback* and *lowfront* ($p = 0.07$).

Results from the regression model shown above suggest that the manipulated stimuli were successful in altering listeners' perceptions of the tokens. *Highback* and *lowfront* were both more likely to be rated as Californian than *highfront* tokens, indicating that neutralizing low and retracted realizations of the short front vowels in both formants had an overall greater impact than neutralizing a single formant. This finding also suggests that low and retracted realizations of the short front vowels in pre-obstruent position are perceptual markers of Californianness (but see §5.3). That *lowback* tokens (manipulated speakers from Hawai‘i) were much less likely to be rated as Californian suggests that speakers may have been attenuating to something else in the speech stream when making their decisions (see §5.1). As the manipulations motivated listeners to make different decisions about the stimuli, the following sections investigate how these different manipulations affected the perception within each vowel category: KIT, DRESS, and TRAP.

4.1 RELATIONSHIPS BETWEEN THE FORMANTS WITHIN VOWELS. The direction of vowel manipulation within each vowel type was also tested to determine whether particular manipulations were more likely to be evaluated as Californian and thus linked more strongly with a perception of Californianness. Separate mixed effects models were fit to binary responses (Californian or non-Californian) for each of the vowels KIT, DRESS, and TRAP. Subject and item were included as random effects, and the manipulation direction (*highfront*, *highback*, *lowfront*, or *lowback*) was included as a fixed effect. All following models were fit to all critical tokens of all vowel types (i.e., all manipulated tokens). An interaction between manipulation direction and gender was also tested for each vowel, as well as an interaction between time spent living in Hawai‘i and manipulation direction. Neither of these interactions yielded significant results.

4.2 KIT. Table 4 shows the results from the model with manipulation direction as a fixed effect and speaker and item as random effects.

Table 4: Logistic regression model of KIT responses to all manipulation directions

KIT	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	0.4071	0.1795	2.268	0.0233
Movement.*highback*	-0.1408	0.1897	-0.742	0.458
Movement.*highfront*	0.6884	0.3271	2.104	0.0353
Movement.*lowback*	-1.3024	0.3192	-4.081	< 0.0001

These results demonstrate that for KIT there is no significant difference between the *lowfront* and *highback* guises; in other words, manipulating just F_1 or just F_2 had no effect for KIT. However, listeners were much less likely ($p < 0.0001$) to evaluate *lowback* tokens (i.e., manipulated Hawai‘i speakers) as Californian, suggesting that speakers may have been attending to phonetic variables besides just formants in the speech stream when making their decisions. A rather unexpected result in table 4 is that listeners were more likely to evaluate a voice as Californian if that voice had both formants manipulated away from a canonical CalE position (*highfront* category). As will be discussed, KIT does not behave like the other vowel types in these data. Separate models were run within gender and found that both male and female listeners demonstrated this same trend, though significance was returned only within males ($p = 0.01$). Explanations for these effects will be addressed in §5.

4.3 DRESS. Table 5 shows the results from the model fit to responses to the DRESS stimuli with manipulation direction as a fixed effect and speaker and item as random effects.

TABLE 5: Logistic regression model of DRESS responses to all manipulation directions

DRESS	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	1.3250	0.1977	6.701	< 0.0001
Movement.*highback*	-0.4459	0.2103	-2.120	0.034
Movement.*highfront*	-1.1094	0.3063	-3.622	0.0003
Movement.*lowback*	-1.5389	0.3071	-5.012	< 0.0001

Similar to the trend found in KIT, tokens in the *lowback* guise were also significantly less likely to be evaluated as Californian ($p < 0.0001$). However, vowels in the *highfront* guise were less likely to be evaluated as Californian ($p < 0.001$). Additionally and most interestingly, DRESS in the *highback* guise was less likely to be evaluated as Californian than DRESS in the *lowfront* guise ($p = 0.03$), as indicated by the negative coefficient for *highback*. In other words, when DRESS retained its low quality but was fronted, it was more likely to be identified as Californian than when it retained its backness but was raised. This model suggests that vowel height in the word DRESS appears to be more closely linked with social perceptions of Californianness than vowel backness. That the *lowfront* guise was perceived as more CalE than the *highback* guise for DRESS will be discussed further in §5.

4.4 TRAP. Table 6 shows the results from the model fit to responses to the TRAP stimuli with manipulation direction as a fixed effect and speaker and item as random effects.

Table 6: Logistic regression model of TRAP responses to all manipulation directions

TRAP	Estimate	Std. Error	z value	Pr (>|z|)
(Intercept)	1.40012	0.17471	8.014	< 0.0001
Movement.*highback*	-0.04813	0.22351	-0.215	0.8295
Movement.*highfront*	-1.19732	0.30260	-3.957	< 0.0001
Movement.*lowback*	-1.12990	0.30348	-3.723	0.0002

Similar to the results found with DRESS, vowels in the *lowback* (p = 0.0001) and *highfront* (p < 0.0001) guises were both significantly less likely to be evaluated as Californian than the *lowfront* guise. Like KIT, no significant results were returned between the *lowfront* and *highback* guises, indicating that manipulating individual formants had no consistent effect on listeners' perceptions.

5. DISCUSSION

5.1 PERCEPTUAL STATUS OF THE SHORT FRONT VOWELS IN CALIFORNIA. Results from this experiment corroborate evidence that the relative positions of formants are important for dialect identification (Baker, Eddington, and Nay 2009) and that social and phonetic information is linked in the mind (Niedzielski 1999; Clopper and Pisoni 2004). Vowels in the *highfront* guise were designed to determine whether vowels could still be identified as Californian in the absence of the formant cues that phonetically characterize the dialect. The results show that with the notable exception of KIT (to be discussed in §5.3), listeners were less likely to evaluate these tokens as Californian. Listeners also rated Hawai'i speakers as "Non Californian," even when manipulated in the direction of the CalE shift. This result lends itself to two possible interpretations, neither of which is mutually exclusive: either (1) there was an additional phonetic cue in the signal that listeners were attending to, and/or (2) the formants were manipulated beyond observed CalE F_1 and F_2 values, and thus were not perceived as Californian. The second of these two possibilities is aided by recent evidence from Grama et al. 2012 that the short front vowels in Hawai'i English are lowered and retracted in pre-obstruent position, likely meaning the vowels were overly shifted (see figure 3).

This experiment also shows that low, retracted realizations of the short front vowels TRAP and DRESS in pre-obstruent position are perceptually relevant markers of CalE. For these two vowels, tokens that were manipulated to be less Californian (those in the *highfront* guise) were significantly less likely to be evaluated as Californian than either of the manipulations in only one dimension. Therefore, *highfront* realizations of DRESS and TRAP in pre-obstruent positions are perceived as non-CalE in this study.

5.2 THE SOCIAL PRE-EMINENCE OF F_2 REVISITED. The results show that for all tokens, both F_1 and F_2 seem to be important when identifying dialect. However, responses to DRESS tokens provide clear evidence that listeners are using vowel height more than vowel backness when identifying a speaker as Californian for this vowel. This result runs contra to the claim made in Labov (2001:167–68) that "English speech communities appear to use … differences in F_2 for establishing social identity." In the current experiment, F_1 is not only used for categorical differentiation but also plays significantly into the percept that a speaker is Californian. Furthermore, the results for TRAP suggest that both formants are important for identifying a vowel as CalE, as neither the *highback* nor the *lowfront* guise was more reliably heard as CalE. Together, these results suggests that both F_1 and F_2 are important acoustic correlates to consider when discussing ways in which social variation may manifest in both production and perception, and that, in fact, F_1 may sometimes play an even more important role than F_2.

5.3 THE PUZZLE OF KIT. The story with KIT is, however, quite a bit different. As noted, KIT was the only short front vowel to be evaluated as more Californian in the *highfront* guise, and it was overall least likely to be evaluated as Californian regardless of the manipulation direction. It is possible that this result is an artifact of the methodology, as KIT was manipulated with respect to reference vowels outside of the short front chain series (i.e., FLEECE and FACE). However, the amount KIT was manipulated was similar in Hertz to that of DRESS and TRAP (see figure 3), a fact that suggests that either (1) speakers used to make the stimuli did not exhibit a shifted KIT or (2) KIT does not manifest in production the same way that DRESS and TRAP do, and listeners are sensitive to this difference. The latter answer is supported by Kennedy and Grama (2012:49), who demonstrate that KIT is the least obviously shifted of the short front vowels in production for a sample of native Southern Californians. While these two possibilities cannot be addressed further with the current experiment, there is sufficient evidence to claim that a lowered and retracted KIT is not as perceptually important to the identification of CalE as either DRESS or TRAP.

6. CONCLUSION. In sum, this study demonstrates that for CalE speakers evaluating CalE voices, the short front vowels DRESS and TRAP are indexical of Californianness. Additionally, results for DRESS tokens suggest that F_1 is more reliably correlated with the percept of Californianness than F_2. This finding runs counter to the claim made by Labov (2001:167) that the second formant is used chiefly to establish social identity. Instead, social meaning can lie in F_1 (as with DRESS tokens in CalE) or both F_1 and F_2 (as with TRAP tokens in CalE). Despite this, it is also clear that formants are far from the only factors in the speech stream that influence listener perceptions. Despite having been manipulated in both F_1 and F_2, speakers from Hawai‘i were much less likely to be perceived as Californian. Furthermore, it is not clear what listeners were attending to when evaluating KIT tokens, as this vowel did not behave like the other short front vowels. Though further experimentation is necessary to tease apart what phonetic variables influence the perception of KIT in California, the results from DRESS and TRAP suggest that both of the major formants that characterize vowel variation should be considered in socio-perceptual and sociophonetic research.

REFERENCES

BAKER, WENDY; DAVID EDDINGTON; and LINDSEY NAY. 2009. Dialect identification: The effects of region of origin and amount of experience. *American Speech* 84:48–71.

CAMPBELL-KIBLER, KATHERINE. 2007. Accent, (ING), and the social logic of listener perceptions. American Speech 82:32–64.

CLARKE, SANDRA; FORD ELMS; and AMANI YOUSSEF. 1995. The third dialect of English: Some Canadian evidence. *Language Variation and Change* 7:209–28.

CLOPPER, CYNTHIA G., and DAVID B. PISONI. 2004. Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change* 16:31–48.

DRAGER, KATIE; M. JOELLE KIRTLEY; JAMES GRAMA; SEAN SIMPSON; and REBECCA CLIFFORD. 2012. Phonetic variation and change in Hawai‘i English. University of Hawai‘i at Mānoa Linguistics Tuesday Seminar. May, 2012.

ECKERT, PENNY. 2008. Where do ethnolects stop? *International Journal of Bilingualism* 12:25–42.

FRIDLAND, VALARIE, and SOHEI OKAMOTO. 2009. The influence of regional vowel shifts on front vowel perception. Paper presented at NWAV 38. University of Ottawa, Canada.

FRY, D. B.; ARTHUR S. ABRAMSON; PETER D. EIMAS; and ALVIN M. LIBERMAN. 1962. The identification and discrimination of synthetic vowels. *Language and Speech* 5:171–89.

GORDON, MATTHEW J. 2006. The West and Midwest: Phonology. In *A handbook of varieties of* English, ed. by Bernd Kortmann and Edgar W. Scheider, 338–50. New York: Walter de Gruyter.

GRAMA, JAMES; M. JOELLE KIRTLEY; KATIE DRAGER; SEAN SIMPSON; and REBECCA CLIFFORD. 2012. Language change in Hawai‘i: KIT, DRESS, and TRAP in motion. Paper presented at New Ways of Analyzing Variation 41, Bloomington IN.

GRAMA, JAMES, and ROBERT KENNEDY. 2009. Acoustic analysis of California vowels. Poster presented at the 83rd Annual Meeting of the Linguistic Society of America, San Francisco CA.

HAGIWARA, ROBERT. 2005. Monophthongs and formant movement in North American English. Poster presented at the 150th meeting of the Acoustical Society of America, Minneapolis MN.

HAY, JENNIFER, and KATIE DRAGER. 2010. Stuffed toys and speech perception. *Linguistics* 48:865–92.

HAY, JENNIFER, PAUL WARREN, and KATIE DRAGER. 2006. Factors influencing speech production in the context of a merger-in-progress. *Journal of Phonetics* 34:458–84.

HINTON, LEANNE; BIRCH MOONWOMON; SUE BREMNER; HERB LUTHIN; MARY VAN CLAY; JEAN LERNER; and HAZEL CORCORAN. 1987. It's not just the Valley Girls: A study of California English. *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 117–28.

JOHNSON, KEITH. 1990. The role of perceived speaker identity in F0 normalization of vowels. *Journal of the Acoustical Society of America* 88:642–54.

KENNEDY, ROBERT, and JAMES GRAMA. 2012. Chain shifting and centralization in California vowels: An acoustic analysis. *American Speech* 87:40–57.

KIRTLEY, M. JOELLE. 2011. Speech in the U.S. military: A sociophonetic perception approach to identity and meaning. University of Hawai'i at Mānoa MA thesis.

LABOV, WILLIAM. 2002. Driving forces in linguistic change. Paper presented at the *International Conference on Korean Linguistics*. At http://www.ling.upenn.edu/~wlabov/Papers/DFLC.htm, accessed 15 November 2012.

LABOV, WILLIAM. 2001. Principles of linguistic change, volume 2: Social factors. Malden, MA: Wiley-Blackwell.

LABOV, WILLIAM. 1994. Principles of linguistic change, volume 1: Internal factors. Malden, MA: Blackwell.

LABOV, WILLIAM; MALCAH YAEGER; and RICHARD STEINER. 1972. A quantitative study of sound change in progress: Report on National Science Foundation Contract GS-3287. Philadelphia PA: U.S. Regional Survey.

LADEFOGED, PETER, and D. E. BROADBENT. 1957. Information conveyed by vowels. *The Journal of the Acoustical Society of America* 29:98–104.

LEVON, EREZ. 2007. Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society* 36:533–54.

NIEDZIELSKI, NANCY. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85.

PLICHTA, BARTEK. 2010. http://bartus.org/akustyk/index.php.

PURNELL, THOMAS; WILLIAM IDSARDI; and JOHN BAUGH. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18:10–30.

R DEVELOPMENT CORE TEAM. 2004. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

RAKERD, BRAD, and BARTŁOMIEJ PLICHTA. 2010. More on perceptions of /ɑ/-fronting. *American Speech* 85:431–49.

SCHARINGER, MATHIAS, and ADITI LAHIRI. 2010. Height differences in English dialects: Consequences for processing and representation. *Language and Speech* 53:245–72.

TRAGER, GEORGE L., and BERNARD BLOCH. 1941. The syllabic phonemes of English. *Language* 17:223-46.

VAN DOMMELEN, WIM, and B. H. MOXNESS. 1995. Acoustic parameters in speaker height and weight identification: Sex-specific behavior. *Language and Speech* 38:267–87.

WALLEY, AMANDA C., and JAMES E. FLEGE. 1999. Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics* 29:307–32.

WELLS, JOHN C. 1982. *Accents of English.* Cambridge: Cambridge University Press.

jgrama@hawaii.edu