University of Pennsylvania Working Papers in Linguistics

Volume 24
Issue 2 *Selected Papers from New Ways of Analyzing Variation (NWAV 46)*
Article 18

10-15-2018

Progressive outliers in listener perception of sound change

Sayako Uehara
Michigan State University

Suzanne Evans Wagner
Michigan State University

Progressive outliers in listener perception of sound change

Abstract
This paper explores how vocalic outliers are perceived in the context of an ongoing sound change, with the prior understanding that listeners integrate vocalic outliers in establishing speakers' intended vowel output, but they do so with a bias (Labov, Baranowski, and Dinkin 2010). We focus on the short-e (DRESS) vowel that is undergoing generational change in lower Michigan. We conducted an experiment to determine whether local community members preferentially incorporate progressive outliers (the outliers in the direction of change) over other outliers into their perception of the vowel's mean position. The results indicate that vocalic outliers were integrated into mean calculation of speakers output, concurrent with the prior study (Labov, Baranowski, and Dinkin 2010). Listeners also displayed bias, as the low short-e progressive outliers were accorded more weight when compared with the high short-e outliers that are in the opposite direction of change. Yet this preference was almost exclusively displayed by the female listeners, who also produced a more lowered short-e in a reading task. This suggests that vocalic outliers are important indicators of the direction of sound change in progress, and that women show special perceptual sensitivity to diachronically progressive outliers.

Progressive outliers in listener perception of sound change

Sayako Uehara and Suzanne Evans Wagner*

1 Introduction

Speakers sometimes deviate phonetically from the speech community's (or their own) usual phonemic target. In Figure 1, for example, each diamond symbol is a token of the back rounded vowel /o/ GOAT (notated in the figure as 'owC') in a sample of spontaneous speech from a single individual. The mean position of the nucleus of the phoneme is shown, with the cross-bars indicating one standard deviation from the mean. Most tokens fall within this range. But the tokens on the far outer periphery of the /o/ token cloud are more than two standard deviations from the mean. Labov, Baranowski, and Dinkin (2010:176) define these tokens as outliers. In the present study, we adopt this same definition.[1] Building on Labov, Baranowski, and Dinkin's (2010) work, we ask whether outliers in the direction of sound change have special properties for the listener.

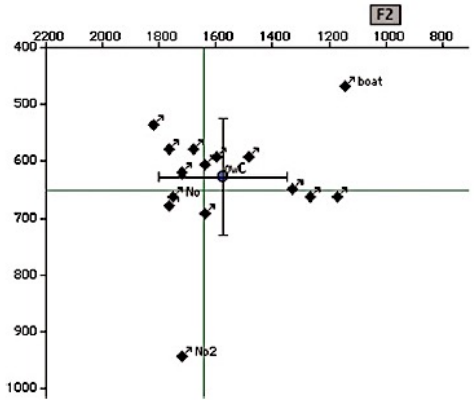

Figure 1: GOAT vowel distribution of a Charleston, South Carolina speaker. Adapted from Labov et al. (2010:176, Figure 2).

Outliers present a challenge to the analyst, who might decide that outliers are unrepresentative of the speaker's output, and exclude them from the linguistic analysis (Ogden 2012). But does this reflect how listeners treat outliers in real life? Labov et al. (2010), whose study we describe in a later section, found that listener rejection or retention of a vocalic outlier token depends on the social information that the outlier conveys. In the current study, we focus on perception of outlier tokens of short-e /ɛ/ DRESS as in *bet, head*. This particular vowel is undergoing generational change in Michigan. Hence we ask whether vocalic outliers in the direction of change, (henceforth 'progressive outliers'), are more socially prominent than other outliers. That is, we seek to determine whether progressive outliers are preferentially integrated into listeners' perception of the vowel's mean position.

2 Outliers and Sound Change

2.1 Outliers as Stylistic Resources

Outlier tokens can be rich sociolinguistic resources for both speaker and listener. Van Hofwegen (2013, 2014, 2015) shows that outliers play an important role in the construction of speaker style,

*We would like to thank MSU Sociolinguistics Lab, MSU Phonology and Phonetics Group, Karthik Durvasula, Aaron Dinkin, Emily Skupin, and the audience at NWAV 46.

[1]Competing statistical definitions of 'outlier' exist (e.g., Leys et al. 2013 cited in Van Hofwegen 2014), but we set them aside here.

stance, and affect. She demonstrates that the outside edges of a vowel distribution are where tokens can be recruited for *stylization*, or sociolinguistic meaning-making (Coupland 2007). Clustering at the edges gives rise to distinctive multimodal token distributions.

Van Hofwegen (2014) also found that vowels undergoing change in progress were more likely than diachronically stable vowels to exhibit these multi-modal distributions and stylized clusters. In her comprehensive study of a single Californian, the speaker's high back vowels /u/ and /o/ GOOSE and GOAT, which are stable in the community, did not exhibit multi-modal distributions, whereas vowels undergoing the California Shift did. The findings, she notes, are in line with those of Eckert (2010) and Podesva (2011), suggesting that there may be a robust correlation between stylized outliers and sound change.[2] This relationship is worth our continued research attention. At the least, it gives rise to the hypothesis that listeners will generally not discard outliers, but recognize them as part of the speaker's distribution for that vowel, doing special sociolinguistic work that in turn may advance sound change.

2.2 Outliers and Sound Change

Labov et al. (2010:177-178) hypothesize a way in which sound change and outlier tokens may interact. The left diagram of Figure 2 shows an outlier token of short-o /ɑ/ as in *pot* (notated in the figure as /o/) falling into the distribution of short-a /æ/ as in *pat*. The other phoneme represented, notated as /oh/, is /ɔ/ as in *caught.* The short-o outlier will either be discarded as an error by the listener, or in some cases, misunderstood as a short-a word. The perceptual mean of the short-o phoneme remains unperturbed, at 1550 Hz F2.

Figure 2: Effect of an outlier with symmetrical distribution (left) and asymmetrical distribution (right) of neighbors. Adapted from Labov et al. (2010:177-178, Figures 3 and 4).

In the right diagram of Figure 2, however, the short-a phoneme has moved forward and up, creating more phonemic distance from short-o. This time, the outlier is less likely to be mistaken for short-a, and may be identified as a member of the short-o class. As the listener integrates this outlier into his or her mean calculation for short-o, the mean will be pulled forward to 1571 Hz F2.

If we combine this scheme with Van Hofwegen's study of stylization, it follows that listeners may pay extra attention to outliers because of their sociolinguistic prominence, which creates a feedback loop that shifts the community's perceptual and production means in the direction of diachronic change.

2.3 Labov, Baranowski, and Dinkin 2010

As mentioned earlier, our experiment is based on a study by Labov, Baranowski, and Dinkin (2010). They investigated a) whether listeners integrate vocalic outliers into their calculation of the mean position of the speaker's phoneme. They were also interested in knowing b) if listeners accord more weight to outliers in the direction of ongoing change when integrating outliers into their calculation.

They focused their study on short-a /æ/ as in *bat*. In the mid-Atlantic area (e.g., Philadelphia, New York City), the nucleus of short-a in this word is often tensed, i.e. raised and fronted to e.g. [ɛæ] under specific phonological, morphological and/or lexical conditions. The word *bad,* produced

[2]Van Hofwegen (2014:31) is careful to underscore that the as-yet limited evidence suggests correlation, not causation, saying we "cannot as yet answer the chicken/egg question: what comes first, the stylization or the linguistic change?"

by a female Philadelphia speaker, was used throughout the experiment. Listeners heard mini token clouds of 5 *bad* tokens. One cloud included an extremely tense *High outlier* token, and one included an extremely lax *Low outlier* token (the others did not have outliers). Upon running a set of perception tasks, the authors found that in general, listeners integrated the outliers into their mean calculations, so that their perceptual mean for short-a shifted accordingly (cf., Figure 2). However, the low outlier had a greater effect than the high outlier on shifting listeners' means. Labov and colleagues proposed that the tense short-a's strong social stigma caused listeners to perceptually correct toward the less stigmatized lax realization.

However, their second question about outliers in the direction of ongoing change remained unanswered, since at the time of the study, tensing of short-a in Philadelphia was considered a nearly-completed diachronic change (Labov 2010). Yet since then, the traditional Philadelphia short-a system has been giving way to a pattern in which tensing occurs only before nasal consonants, as in *ban, Sam* (Labov et al. 2016) but crucially, not *bad*. This means we can entertain a new possibility: That listeners in this study were giving more weight to the low outlier because lax *bad* was an avatar of ongoing sound change, at least for some listeners. Yet this cannot be untangled from mid-Atlantic participants' negative evaluation of short-a tensing more generally.

3 Research Questions

To better address the question of whether outliers in the direction of sound change are given special weight by listeners as they calculate vocalic means, it would be helpful to know what listeners do if a vowel is not a highly socially salient stereotype like tense short-a. Second, it would be useful to examine a vowel that is undergoing more clearly vigorous change in apparent time. Thus, we repeat Labov et al.'s (2010) study, but with a focus on a vowel that is undergoing change below the level of consciousness. We will address these questions below:

a) Are outliers integrated into listeners' mean calculations?
b) Will the vocalic outlier in the direction of change be weighted more heavily than others in calculation of the mean?
c) What can production tell us about perception? Is there any parallelism?

For questions (a) and (b), we ask whether Labov et al.'s (2010) results can be replicated with a different vowel. Question (c) seeks information that Labov et al. (2010) did not have, and may reveal more details about the listeners' perceptual behavior through their production.

4 Short-e Lowering and its Evaluation in the Inland North

The regional locus for our study is the part of Lower Michigan that is located in the Inland North dialect region. The Inland North is defined in the Atlas of North American English (Labov, Ash, and Boberg 2006) by its rotation of six vowels in the Northern Cities Shift (NCS) (Figure 3).

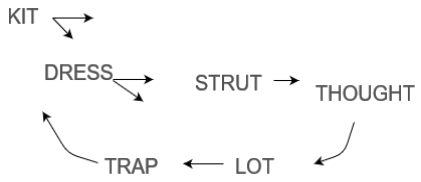

Figure 3: Vowel movements of the Northern Cities Shift.

However, more recent studies of Chicago (McCarthy 2011), Rochester, NY (Maya Abtahian p.c.), Ogdensburg, NY (Thiel 2017) and Detroit (Morgan et al. 2017), among others, suggest that the NCS is no longer in progress. Indeed, some components of the Northern Cities Shift, such as the fronting of LOT, appear to be reversing in some locations. In fact, the NCS may be giving way to a different vowel configuration. It has been observed in so many North American speech communities that Fridland and Kendall (2012) have called it the "Elsewhere Shift." Figure 4 displays the Elsewhere

Shift directions in apparent time for the NCS vowels. Note that in both patterns, DRESS is lowering and backing; henceforward we focus on lowering. Importantly for our study, DRESS-lowering is a robust change in progress across southeast Michigan, as reported by Wagner et al. for Lansing (2016), by Zheng (2018) for Troy and by Morgan et al. (2017) for metro Detroit.

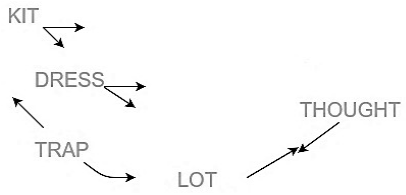

Figure 4. Vowel movements of "Elsewhere Shift."

Equally importantly, this change in progress is happening without the knowledge of the speech community, although in a matched-guise survey of 930 college-aged Lower Michiganders, Savage (2017) found some implicit evaluation of short-e lowering. Young women were more likely than their male peers to perceive a lowered short-e variant as intelligent, confident and articulate, as well as friendly. Furthermore, Zheng (2018) and Nesbitt and Mason (2016) found that young people lowered short-e in formal styles. In sum, DRESS-lowering contrasts well for our purposes with the short-a variable employed in Labov et al.'s (2010) study: It is a clear change in progress, and it is below the level of speaker awareness with no negative evaluation

5 Methods

5.1 Participants

23 white undergraduate students, 13 men and 10 women aged 18 to 24, were recruited from Michigan State University. They were all native speakers of US English from south or southeast Lower Michigan. Their overall normalized vowel system in Figure 5 shows a typical Elsewhere-like split of short-a in pre-nasal and pre-oral contexts (MAN, TRAP), and backed and lowered short-e (DRESS). The mean F2 position of DRESS is in line with Morgan et al.'s (2017) observations in Metro Detroit, where most of the participants grew up. Most importantly, the mean height of DRESS is also in line with the mean for all other Lower Michigan young adults in the aforementioned studies, and none of the individual participants deviate substantially from this mean. Because these participants are similar demographically and linguistically to those in the studies of Lansing and metro Detroit, we have good reason to think that the participant sample is representative of their age group, ethnicity, education level and dialect region.

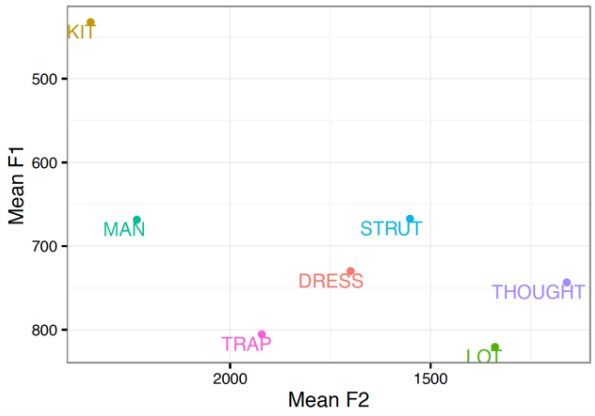

Figure 5: Michiganders' mean vowel space with nasal split for short-a.

5.2 Production Data Collection

The production data above was collected by having participants read a passage, *A Bad Day for Ducks,* created by Dennis Preston and employed in several previous studies of Lower Michigan speech (e.g. Zheng 2018). The passage contains 110 words with NCS vowels in stressed position and twelve of those words contain DRESS. Vowel nuclei were extracted, measured and normalized using FAVE (Rosenfelder et al. 2011).

For all participants, the passage reading was conducted before the perception task in order to avoid priming. Following the passage reading, participants did a perception experiment. This was designed to answer our second research question: "When participants hear extra-low DRESS tokens (*progressive outliers*), which are in the direction of generational change, will they weight them more heavily than other outliers into their calculation of the mean?"

5.3 Perception Data Collection

To test our research questions on perception of progressive outliers, audio stimuli were constructed following the method employed in Labov et al. (2010). A female speaker in her late 20s from the Inland North dialect area was recorded saying *bed*. The nucleus of DRESS for this recording was approximately 800Hz F1 and 1600Hz F2. It was resynthesized in Praat along the F1 continuum of 580~1060 Hz. This gave a range either side of the speaker's natural mean of two standard deviations.

The resynthesized tokens were embedded in 5 carrier phrases (***Stay in*** *bed*, ***It's a great*** *bed*, ***Clean the*** *bed*, ***She's in*** *bed*, ***He makes the*** *bed*) recorded by the same female speaker. None of the vowels implicated in the NCS or Elsewhere Shift appear in these phrases in primary stressed position. We used a within-subject design conducted with four experimental conditions: Low Symmetrical, Low Outlier, High Symmetrical, and High Outlier (Table 1). Each condition consisted of one repetition of each of the five carrier phrases. As shown in Table 1, each condition constituted a mini-cloud of *bed* tokens. In each condition, three *bed* tokens clustered at the center of the distribution. In the Low Symmetrical and High Symmetrical conditions, the other two tokens were placed only one 'step' (40 Hz F1) higher and one step lower than the central cluster. On the other hand, the Low Outlier condition included a token one step higher and one several steps lower than the center, roughly two standard deviations away. Likewise, the High Outlier condition also contained an outlier two standard deviations above the central cluster.

F1 (Hz)	High Outlier	High Symmetrical	Low Outlier	Low Symmetrical
580	**1**			
620				
660				
700				
740		**1**		
780	**3**	**3**		
820	**1**	**1**	**1**	**1**
860			**3**	**3**
900				**1**
940				
980				
1020				
1060			**1**	

Table 1: Distributions of short-e vowel tokens in the four conditions—"High outlier," "High symmetrical," "Low outlier," and "Low symmetrical." The mean values of each conditions were: 748 Hz, 780 Hz, 892 Hz, and 860 Hz accordingly.

5.4 Procedure

The production and perception experiments were presented through the LimeSurvey online platform. Participants completed the tasks in a quiet room, seated in front of an iMac. They recorded their passage reading and listened to the audio stimuli using headsets. During the perception experiment, participants were asked to listen to all of the conditions, one at a time (Low Symmetrical, Low Outlier, High Symmetrical, and High Outlier), and each condition was presented on a separate page. Participants were instructed to listen to the condition on the page as many times as they desired, and to choose a *bed* audio token from {580 Hz, 660 Hz, 740 Hz, 820 Hz, 900 Hz, 980 Hz, 1060 Hz} that best represented the speech in each condition.

6 Results

6.1 Perception Experiment

All the values presented in the perception results were converted into Mel scale.[3] The overall results for the 23 participants are shown in Figure 6.

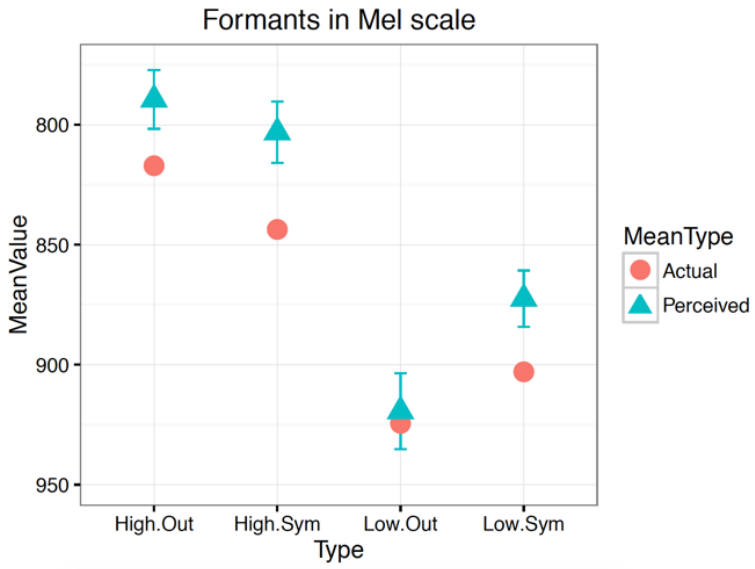

Figure 6: The actual means (red circles) and perceived means (green triangles) of tokens in four conditions: "High Outlier," "High Symmetrical," "Low Outlier," and "Low Symmetrical." The standard error bars are presented for the perceived values and all the mean values are converted into Mels.

The perceived mean values for the High Outlier, High Symmetrical, and Low Symmetrical conditions were all *higher* than the actual mean values. Yet the Low Outlier condition appears to be an odd one out of the four since the mean for this cluster was perceived to be very close to the actual mean value presented in the experiment. From this simple observation, our results seem to indicate two types of bias. The first bias is the consistently high perceptual mean in three conditions: High Outlier, High Symmetrical, and Low Symmetrical. Because of this, it is unclear in the High Outlier condition whether (i) participants' perceptual mean overshot the actual mean because of this apparent general bias, or because (ii) the high outlier token was included to some degree in their calculation. Even if (ii) is the correct interpretation, it is still clear that participants exhibited a second bias. They seem to have integrated the low outlier token into their calculation of the mean to a greater

[3]The mel scale is a scale of pitches judged by listeners to be equal in distance one from another. $m = 2595\log10(1+f/700)$

degree than they did for the high outlier token. This led to greater accuracy in calculation of the mean for the Low Outlier condition than for the High Outlier condition. Recall that the low outlier token is in the direction of generational change, providing preliminary support for the hypothesis that listeners accord more weight to progressive outliers than other types of outliers.

6.2 Perception Results by Gender

To further interpret this perception data, we looked at the data by gender. As Figure 7 shows, symmetrical conditions did not exhibit obvious differences between female and male responses. But in the outlier conditions, it appears that the general bias toward a higher perceptual mean is greater for male participants than female participants. More interestingly, in the Low Outlier condition, men exhibited the bias toward a higher mean, but the female perceptual mean was even lower than the actual mean. Like a subset of individuals in Labov et al.'s (2010) study, the women seem to treat the low outlier as the perceptual target, rather than the cluster around the mean.

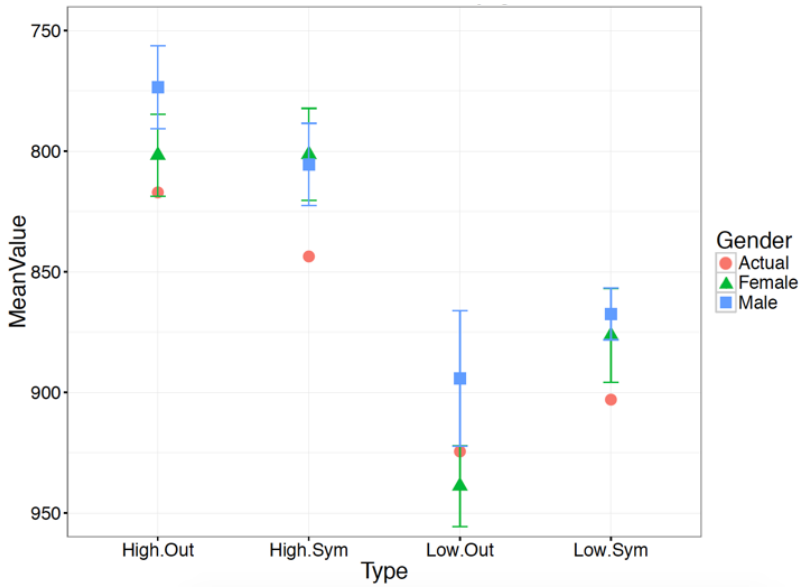

Figure 7: Perception results by gender for each condition. Values are all in Mels.

In an earlier section, we mentioned that young women were more likely than their male peers to perceive a lowered DRESS variant as intelligent, confident, articulate, and friendly (Savage 2017). It may well be the case that female participants were more willing to integrate the low outlier DRESS token than the male participants, because women have a more positive evaluation of low tokens. However, this remains speculation awaiting further investigation.

6.3 Production Results by Gender

A less speculative source of information is the participants' own production of DRESS, from the *Bad Day for Ducks* reading passage. The two plots in Figure 8 give the normalized mean values for male and female participants. For DRESS, women's mean vowel value is lower than men's. The difference between DRESS means for male and female was significant ($w = 19$, $p = 0.003243$). That is, women are more advanced with respect to this change than men. Looking at the DRESS means by individual participant (Figure 9), we can observe that the men and women's means are generally separate, with only three men having means in the same F1 range as the women.

Even though we cannot say that women's low production causes them to give special weight to low outliers, or vice versa, we can say that they may be related. In other words, if it is the case that their production here represents their production prototype for this particular vowel, then their perception prototype is likely to be overlapped with the production prototype for both men and women.

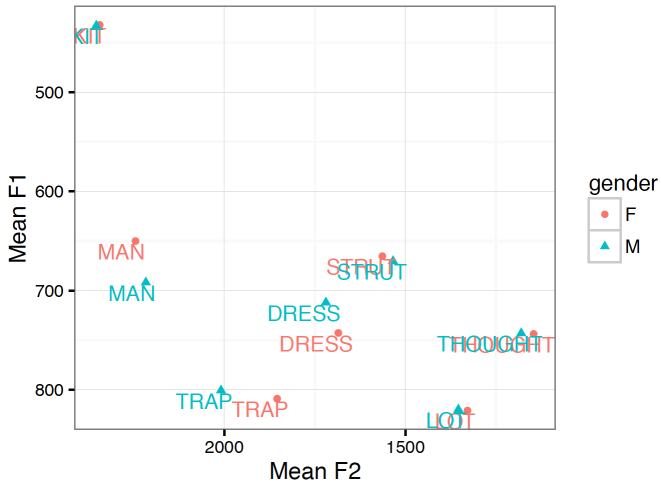

Figure 8: Production results of passage reading for 23 participants. Overall vowel means by gender. Red circles are female, blue triangles are male.

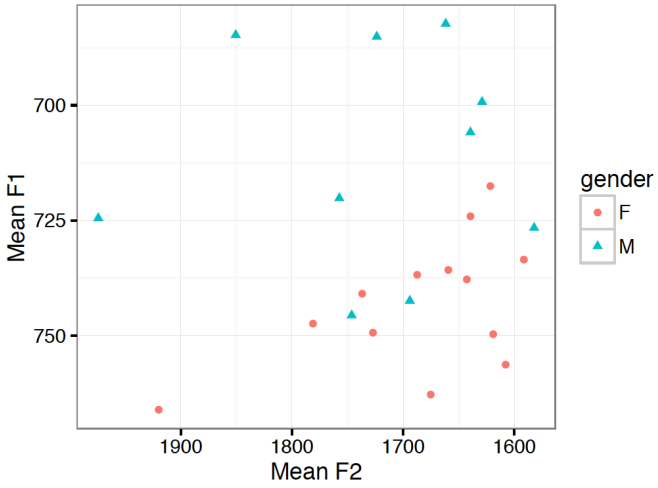

Figure 9: Short-e individual production means of 23 participants from passage reading. Red circles are female, blue triangles are male.

7 Discussion

To return to the three research questions of the study, we first asked, "*Are outliers integrated into listeners' mean calculations*?" Inferring from the results, listeners are generally attentive to vowel tokens that are on the far periphery of phoneme token clouds. Although there was an obvious discrepancy between which outlier was perceived accurately and by the gender of the listener, participants (female) were able to integrate the Low Outlier token into their calculation of the perceptual mean. On the other hand, it is uncertain whether the outlier in the High Outlier condition had a very strong effect, pulling participants' perceptual target past the mean, or whether this was an effect of overall bias toward a high perceptual mean for DRESS that we can also see for the symmetrical conditions. Nevertheless, it is interesting that the High Outlier is behaving similarly to the symmetrical types (High Symmetrical and Low Symmetrical). It is unlikely that the High Outlier in this condition

itself is pulling listeners' perception because it is treated the same way as the two symmetrical conditions. The results are not any higher than High and Low Symmetricals. That being said, it is unclear if the outlier token in the High Outlier condition is ignored by the listener or if that token is actually not an outlier. It might be the case that this token is close to the prototype of DRESS and the perceptual space of DRESS is overlapping, which cause the token to sound normal for the listener and the low outlying token in Low Outlier condition to sound more salient.

Regarding the results for High Outlier, High Symmetrical and Low Symmetrical, it appears that participant accuracy was impeded by some kind of bias toward a high perceptual mean for DRESS. The question is, was this bias socially motivated? If participants had a social preference for a high DRESS that is very conservative with respect to sound change, then this mental prototype may have interfered with their accuracy. This is similar to the effect that Niedzielski (1999) found in her experiments with Michigander participants, where their beliefs about the 'normalness' of Michigan speech affected perceptual accuracy.

In our second research question, we asked, "*Will the vocalic outlier in the direction of change be weighted more heavily than others in calculation of the mean?"* Our findings show that it was women who seemed to treat the Low Outlier as socially salient. Female participants undershot the F1 target mean in the Low Outlier condition compared to other three conditions. It is likely that female Michiganders are paying special attention to outliers in the direction of change. They behave somewhat like the four "outlier individuals" in Labov et al.'s (2010) experiment (interestingly, all men), although these four seemed to treat both low and high outliers "as indicative of the target position of a phoneme" (Labov et al. 2010:188). In this regard, another question arises: *why would women treat the Low Outlier token as a target, not an outlier?* It was mentioned above that the progressive low DRESS was shown in a previous study to be evaluated somewhat positively by young women. Such an evaluation might be enhancing their receptiveness to low DRESS in our experiment.

Finally, our third question asked: "*What can production tell us about perception?*" From our results, it seems that the production of DRESS is related to perception of DRESS. As stated earlier, female participants have a significantly lower DRESS than do male participants. They are more advanced on the path of community sound change, and thus may have more exposure to exemplars that are in the vanguard of that change, which allow them to form more advanced mental prototypes of the central tendency for the phoneme. Accordingly, in our experiment, female participants give the outlier in the direction of change special attention while male participants maintained the high perceptual bias for the Low outlier.

8 Conclusion and Future Questions

To conclude, it is very likely that outliers are meaningful tokens for listeners, as well as for speakers, as other studies have claimed. However, we also need to consider the bias listeners might have toward specific kinds of vocalic outliers. We saw that there was an apparent bias with respect to which outlier had a stronger impact on perception. In our study, it was the low DRESS outlier in the direction of the change – the progressive outlier – that was accorded more weight. Interestingly this outlier is somewhat positively evaluated by Michiganders, concurrent with Labov, et al.'s (2010) findings that a non-socially-stigmatized outlier was accorded more weight in perception. Additionally, female participants produced, in their own speech, a lower overall DRESS mean than male participants. This might have contributed to women participants' special attentiveness towards low outlier token. We are left with an outstanding question: To trigger the effect that we obtained in this study, are all three independent variables required? That is, must the outlier be progressive (in the direction of the change), and be evaluated positively and must the listeners themselves have progressive realizations of the vowel in question?

Moreover, there are still other matters to consider. It would be worth investigating if we see the same effect in perception for older participants. Participants in the current study are college-aged Michiganders. We assume that older speakers in the same speech community are aware, at some level, of the direction of change as it has been happening for decades. Yet, they might not have been exposed to many progressive exemplars that are this extreme compared to young adults; so it would be interesting to see if they would be more likely to reject the outlier than the young speakers. Also, it would be worthwhile studying the effects on other vowels that are undergoing sound change in Michigan. We do not know if the perceptual effect we see in our study is particular to DRESS, and

specifically to short-e lowering in Michigan. Since the Elsewhere Shift is reported to be occurring in other parts of the U.S., it would be interesting to see if the non-Inland North region with ongoing short-e lowering will observe a similar effect as this one.

References

Eckert, Penelope. 2010. Affect, sound symbolism, and variation. *University of Pennsylvania Working Papers in Linguistics* 15.2, ed. by K. Gorman and L. MacKenzie.

Labov, William. 2010. Social Evaluation of the Northern Cities Shift. *Principles of Linguistic Change, Volume 3*, 236–244.

Labov, William, Sharon Ash and Charles Boberg. 2006. *Atlas of North American English: Phonetics, Phonology, and Sound Change.* New York, Berlin: Mouton de Gruyter.

Labov, William, Maciej Baranowski, and Aaron Dinkin. 2010. The effect of outliers on the perception of sound change. *Language Variation and Change* 22:175–190.

Labov, William, Sabriya Fisher, Duna Gylfadottir and Betsy Sneller. 2016. Competing systems in Philadelphia phonology. *Language Variation and Change* 28:273–305.

Leys, Christophe, Christophe Ley, Oliver Klein, Philippe Bernard, and Laurent Licata. 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* 49:764–766.

LimeSurvey Project Team / Carsten Schmitz. 2015. LimeSurvey: An Open Source survey tool. LimeSurvey Project Hamburg, Germany. http://www.limesurvey.org

Morgan, Beau-Kevin, Kelsey DeGuise, Eric K. Acton, Daniele Benson and Alla Shvetsova. 2017. Shifts toward the supra-regional in the Northern Cities region: Evidence from Jewish women in Metro Detroit. New Ways of Analyzing Variation (NWAV46): Madison, Wisconsin, November.

McCarthy, Corrine. 2011. The Northern Cities Shift in Chicago. *Journal of English Linguistics,* 39:166–187.

Nesbitt, Monica. and Alex Mason. 2016. Dialect leveling across the US: Evidence of the Third Dialect Shift in the Inland North. Paper presented at New Ways of Analyzing Variation 45, Vancouver, Canada.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology* 18:62–85.

Ogden, Richard. 2012. Making sense of outliers. *Phonetica* 69:48–67.

Podesva, Robert J. 2011. The California Vowel Shift and gay identity. *American Speech* 86:32–51.

Rosenfelder, Ingrid, Josef Fruehwald, Keelan Evanini, and Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. Available at: http://fave.ling.upenn.edu.

Savage, Matt. 2017. Attitudinal responses to two NCS vowels among Inland North speakers. Comprehensive exam paper, Michigan State University.

Thiel, Anja. 2017. The Northern Cities Shift going Elsewhere: Evidence of the Elsewhere Shift in Northern New York. (Submitted). In *Sociolinguistic Symposium* 2018.

Wagner, Suzanne E., Alex Mason, Monica Nesbitt, Erin Pevan and Matt Savage. 2016. Reversal and re-organization of the Northern Cities Shift in Michigan. In *University of Pennsylvania Working Papers in Linguistics* 22.2 ed. by H. Jeoung.

Van Hofwegen, Janneke. 2013. Vocalic style versus stylization: How outliers exemplify acoustic axes of style. New Ways of Analyzing Variation (NWAV 42): Pittsburgh, PA. October.

Van Hofwegen, Janneke. 2014. The new normal: Multi-modal distributions signifying loci of vocalic stylization. Qualifying Paper, Stanford University, CA.

Van Hofwegen, Janneke. 2015. The new normal: Multi-modal distributions signifying loci of vocalic stylization. Linguistic Society of America 2015 Annual Meeting: Portland, OR. January.

Zheng, Mingzhe. 2018. You have to learn to adapt: A sociolinguistic study of Chinese Americans in the "Asian City" of southeast Michigan. Unpublished doctoral dissertation, Michigan State University.

Department of Linguistics and Germanic, Slavic, Asian and African Languages
Michigan State University
East Lansing, MI 48824-1027
ueharas2@msu.edu
wagnersu@msu.edu