Connecting production and perception on an individual basis: Back vowel fronting in Southeast Wisconsin

Jonathan Jibson
University of Wisconsin–Madison

1 Introduction

A common indication of the state of a sound change in progress is acoustic data elicited in sociolinguistic interviews or production tasks – what people say. Perceptual studies, though, are also integral in that they indicate what forms are salient to speaker-listeners on some level. The present study seeks to connect production and perception data in a new way, considering how individual speakers' perception of a certain form correlates with their own production of that form (rather than aggregating such data). The results suggest that this connection is reasonably clear for regionally affiliated forms but that the strength and direction of this correlation varies from form to form.

Back vowel fronting (BFV) is a change in progress whereby back vowels – in English, the high back vowels /u, ʊ/ and mid-high back vowel /o/ – are articulated closer to the front of the mouth and/or with less lip rounding than in the past. This change is not unique to American English varieties (Kleber, Harrington & Reubold 2011), nor is it unique to the English language (Labov 1994; Harrington 2012).

In North America, BVF has long been associated with Southern varieties of American English (Thomas 1989; Fridland & Bartlett 2006). In discussing the Southern Vowel Shift, Clopper, Pisoni & de Jong (2006:1662) say, "The primary attribute of this shift is the fronting of the back vowels /u/ and /o/." The data they present also show fronting of /ʊ/.

However, in American English, this feature is entering other varieties as a sound change. Fridland & Bartlett (2006) cite research on BVF in the Midwest (Ash 1996), the West (Hagiwara 1997), and other areas. The main pattern is that South varieties' back vowels are most fronted and that non-Southern varieties are fronting

/u/ more than /o/. Baronowski (2008:528), in discussing this general spread, adds this caveat: “Although there have been reports of /uw/ fronting in the North … the dialect remains conservative in the fronting of /ow/.” In this way, BVF marks different regionally affiliated dialects differently: Southern varieties front the higher back vowels the most, the Midlands and the Western varieties front them but to a lesser extent, and (Inland) Northern varieties front /u, ʊ/ the least and do not front /o/. What we do not know, however, is whether speakers are actually aware of these differences.

Perceptual data can enrich acoustical findings. That is, in studying speech, assessing what people say is completed by assessing what people hear. In Plichta & Preston (2005), an audio recording of the word *guide* was resynthesized into progressively more monophthongal steps from a fully diphthongal one, and participants were asked to place each audio clip on a north-to-south map with nine potential cities. In the end, participants rated the fully diphthongal step as most northern and each successive step as more southern without exception. Here, when speakers knew what the word was, they relied on their belief that monophthongal /ai/ is a Southern American English feature and responded as such.

Fridland, Bartlett & Kreuz (2004:4,13) discuss the importance of connecting production and perception data. They articulate succinctly the question as “whether speakers are more perceptually aware of changes they themselves take part in.” Their findings indicate the following as an answer to that question:

> The ability of participants to accurately rate differences between vowel variants and assign scores appears to vary, depending on whether the local community speech norms involve those particular variants and whether those variants are shared with other regions (e.g., back vowel fronting). … So, the shifts that affect Memphians in particular, not only Southerners in general, appear to be most salient perceptually to listeners [in Memphis].

The connection between production and perception has largely been sought out by comparing the results of related but disparate studies. Perceptual studies like that of Fridland, Bartlett & Kreuz (2004) draw on the solid literature on the Northern Cities and Southern vowel shifts for production data (see the first paragraph of their article for references). This approach typically aggregatesand generalizes results from large numbers of participants. The connection between one person’s production and perception – and the difference between that and the connection between another person’s production and perception – can be obscured in the process. Harrington, Kleber & Reubold (2008) come closer to investigating the individual connection, comparing the production and perception of back vowels for the same group of individuals, but the results are reported only in averages and not on a person-by-person basis.

Natvig (2014) explored /æ/ raising experimentally in order to understand how one's productin of this American regionalism relates to one's perception of it. Preapical and prevelar /æ/ tokens were elicited from participants, after which participants judged the same tokens recorded from others of varying raising patterns as being from the North (with prevelar raising), the Northern Cities (with preapical raising), or the Midlands area (with no raising). He found that participants with little versus much raising in their production tended to pattern alike in their perception judgments. However, these results came from grouping participants into "raisers" and "non-raisers," a decision that was intuitive and obvious to make after the production data came in (Natvig, personal communication, 23 December 2016). Although this is more individualized than any other data I have yet found, it still does not treat every individual differently in terms of their gradient production.

With these considerations, I come to my basic research questions. Will I find the same acoustic results from production data with respect to BVF that have been established in the literature? Do Wisconsin speakers use BVF as a regional marker? Does the answer to the second question depend on the answer to the first question? Based on these questions, I hypothesize as follows:

1. The group of participants will perceive changes better in *hoed* than *hood*/*who'd*.
2. The group of participants will front *who'd* and *hood* somewhat but not *hoed*.
3. Individual participants producing a more fronted back vowel will better perceive differences in F2 for that vowel than individual participants with less fronted back vowels will.

2 Methodology

2.1 Stimuli

The audio stimuli were collected first. These came from five female speakers born and raised near Green Bay, Wisconsin and one female speaker born and raised near Raleigh, North Carolina. All speakers were personal acquaintances. These speakers were recorded reading a text that elicited three tokens each of *who'd, neighborhood*, and *Hode/hoed*. This text is included as Appendix A. The use of *neighborhood* instead of *hood* was simply a constraint of writing a sensible narrative. A narrative was used instead of a word-in-frame reading list because the former was a more natural task for participants and thus made them more amenable to participating.

The formant values of F1, F2, and F3 were measured for the *hoed*, *hood*, and *who'd* tokens, and a separate audio clip was made for each token. One speaker, given the label WI, was selected for analysis, and one token each from *hoed*, *hood*, and *who'd* were used as test items in the present study. Likewise, the most natural sounding token of each vowel type was selected for the North Carolina speaker, who has been given the label NC.

The tokens from WI and NC were then resynthesized using Praat. Each recording was resampled down to 11 kHz. For each target word, four new audio clips were created, each progressively 100 Hz different in F2 from the previous step. For WI's tokens, these Hz values were successively added (fronting); for NC's tokens, these Hz values were successively subtracted (backing). The result was a set of five audio clips for each of the three vowels for both of the two speakers, for 30 test items total.

After the test items were prepared, 50 distractor items were created from the same speakers and from others. In the end, 80 stimulus items were created.

2.2 Experiment

I recruited five participants who were college-aged, female, born-and rasied Wisconsinites. Young females are canonically at the forefront of sound change (Labov 1994; Clopper, Pisoni & de Jong 2006; Jacewicz, Fox & Salmons 2011), and so I hoped this group would show BVF.

The production portion of the experiment was carried out first. Participants were audio-recorded reading a slightly revised version of the aforementioned narrative reading (see Appendix B).

The perception portion of the experiment was carried out through Qualtrics. First in the survey was a basic demographic information request. Next, the audio stimuli were presented, with one item on the screen at a time. The order was randomized for each participant. Participants were instructed to rate each recording as *definitely Wisconsin, probably Wisconsin, probably not Wisconsin*, or *definitely not Wisconsin*.

2.3 Analysis

Acoustic analysis of the narrative reading involved forced alignment with FAVE (Rosenfelder et al. 2011), after which I checked all relevant boundaries by hand. A Praat script gathered the formant values for F1, F2, and F3 from the head and tail of each vowel, defined as at least three pulses each at 30% and 70% time points unless poor formant tracking necessitated moving these boundaries (which was seldom the case) either earlier or later in the vowel, depending on which was closer to the visible formants in the poorly tracking section. Hertz values were also converted into Bark and normalized using the Bark Difference metric. For each speaker, all tokens of a vowel were averaged.

An analysis of the perceptual experiment focused on plotting responses to related audio clips on a chart, fitting a trend line, and calculating the slope and R^2

coefficient of that trend line. In this way, I could ask whether, say, a more fronted or more backed *hoed* vowel was judged as being more or less Wisconsin sounding for both individual participants and the entire group. Two other scores were also used: QUAL and AUG. Both involved averaging the judgments to one audio clip from WI to those of a corresponding clip from NC. For QUAL, the unmodified WI clip was connected with the most modified NC clip, since these were each the most representative of Wisconsin English of its continuum; For AUG, the unmodified clip from each was the focus, as was the most modified clip from each. In this way, I could ask whether formant values' representativeness of Wisconsin English was more or less important to participants than the quality of the audio file itself.

Finally, the perceptual data was treated as a dependent variable of the acoustic data. A bar chart was created for each word (*hoed, hood,* and *who'd*). The participants were ranked along the X-axis according to their representativeness for that feature, following the non-normalized speaker-internal method outlined in Evans & Preston (2001). The Y-axis consisted of R^2 values of the trend line from the QUAL continuum for the word in question. Another set of bar charts was created in which the X-axis was the same but the Y-axis was the slope of the trend lines rather than the R^2 values.

3 Results

3.1 Production

Figure 1 below shows the average acoustic values for each speaker for *hoed, hood,* and *who'd*. (The *head* tokens are included because they were used as speaker-internal anchor points for further analysis, as described below.) As expected, the /u/ vowels showed the greatest variation across speakers, with the /o/ and /ʊ/ vowels being comparably consistent across speakers. It would appear, based on these findings, that /u/ is relatively variable while /o, ʊ/ are relatively consistent. Looking at the Parents and Grandparents generations reported in Jacewicz, Fox & Salmons (2011), we see that /u/ and even /ʊ/ are considerably more fronted than is traditional in Wisconsin – by about 300 and 400 Hz, respectively – and /o/ is fronted by about 100 Hz.

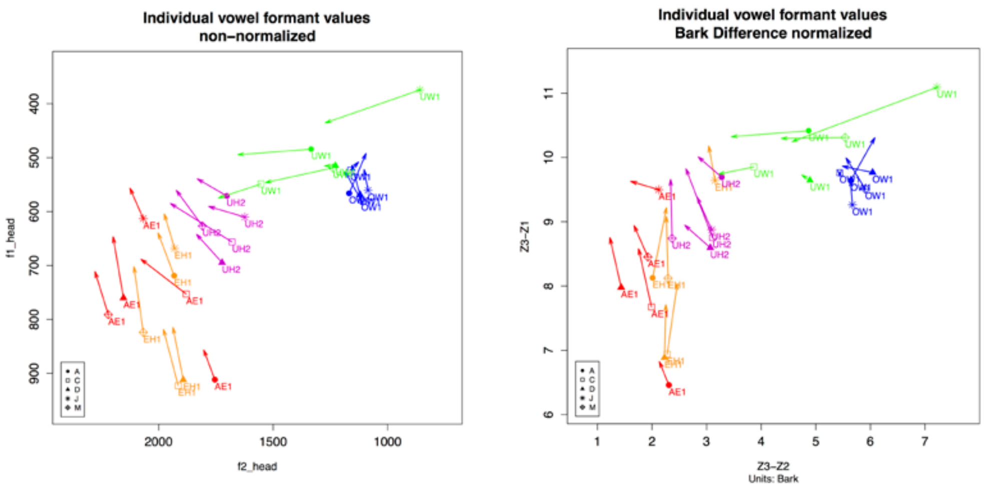

Figure 1: Speaker average productions, with Hertz values on the left and Bark Difference values on the right

To situate these results in a larger context, the group's average Hertz values for F2 can be compared to those of other studies. Clopper, Pisoni & de Jong (2006) included data on the North dialect area. Their measurements were taken at one-third time points, which correspond closely to my 30% time point measurements. My /ʊ/ values were more fronted, with my ellipsis only overlapping theirs by a small margin. My /o/ values were well within their ellipsis but clearly less fronted than the their main cluster of tokens. My /u/ values were also well within their ellipsis, though their group is heavily skewed by only a small number of very fronted outliers where my data is heavily skewed by one back outlier (and so my data appear much more fronted than theirs).

My data conform more closely with the findings of Jacewicz, Fox & Salmons (2011), as is to be expected given that their participants and mine came more narrowly from Wisconsin and not from a broader sample of the Northern dialect area. Table 1 presents my 30% time point values compared to the 35% time point values from Jacewicz, Fox & Salmons (2011) across generations. My findings are closer to the Girls generation than to P(arents) or G(randparents), which are below dubbed Grandmothers and Mothers. This fits with our expectations, since "Girls" was defined as 8–12 year olds and their study was published in 2011 (so that 18–22 year olds today are from the same age cohort). We can see that *hood* is more fronted in my data than in the Girls data but that *hoed* and *who'd* are less fronted. However, these differences are only about 50 Hz for each.

Table 1: Head F2 values in Hz from three generations in Southeast Wisconsin compared to data from the present study

Word	Grandmothers	Mothers	Girls	J
Hoed	871	973	1170	1118
Hood	1207	1276	1639	1694
Who'd	933	998	1258	1218

I followed Evans and Preston's (2001) speaker-internal method for ranking speakers from most to least fronted. Using *head* as the anchor in this speaker-internal indexing approach, we can rank speakers from least to most fronted /u, ʊ, o/. These rankings are used for the production-perception connection portion of the analysis below. Also, from these rankings, it became clear that speakers tended to be ranked about the same for all back vowels. Table 2 displays the rankings for each vowel as well as the average for the three back vowels, with 1 being the most representative of Wisconsin English in the group.

Table 2: Speaker rankings according to Wisconsin English representativeness

Participant	/u/	/ʊ/	/o/
A	5th	4th	4th
C	6th	6th	5th
D	4th	5th	6th
E	2nd	1st	1st
J	1st	2nd	3rd
M	3rd	3rd	2nd

For a deeper look at the connection between the rankings of one vowel and those of another, I calculated the R^2 coefficients for each pair of vowels in Table 3. What we see is that /ʊ/ has an equal R^2 coefficient for its correlation with /u/ and its correlation with /o/, but /u/ and /o/ have a notably lower value. These results suggest that the forms associated with BFV co-occur to a fair extent.

Table 3: R^2 coefficients for speaker rankings between vowel pairs

	/u/	/ʊ/	/o/
/u/	-	0.78	0.43
/ʊ/	-	-	0.78
/o/	-	-	-

3.2 Perception

As an illustration of the results from the perception experiment, I have displayed in Figure 2 the four graphs for the judgments of *hoed* averaged across all speakers. The top two graphs show judgment scores to WI's and NC's tokens. The bottom two graphs explore whether judgments were more influenced by formant values or by the progressively more artificial sounding quality of the audio: the bottom left graph, QUAL, shows WI's 0 averaged with NC's 4, WI's 1 averaged with NC's 3, and so forth; the bottom right graph, AUG, shows WI's 0 averaged with NC's 0, WI's 1 averaged with NC's 1, and so forth. Judgments had the possible values of 1, 2, 3, or 4, but in these graphs the axis is 1 (definitely not Wisconsin) and the upper boundary is 3 (probably Wisconsin). NC's trend line should have the opposite slope from WI's because step 0 for NC is its least Wisconsin sounding token, not its most Wisconsin sounding token. QUAL should have positive slope values for the same reason that WI should, and I guessed that AUG would have negative slope values.

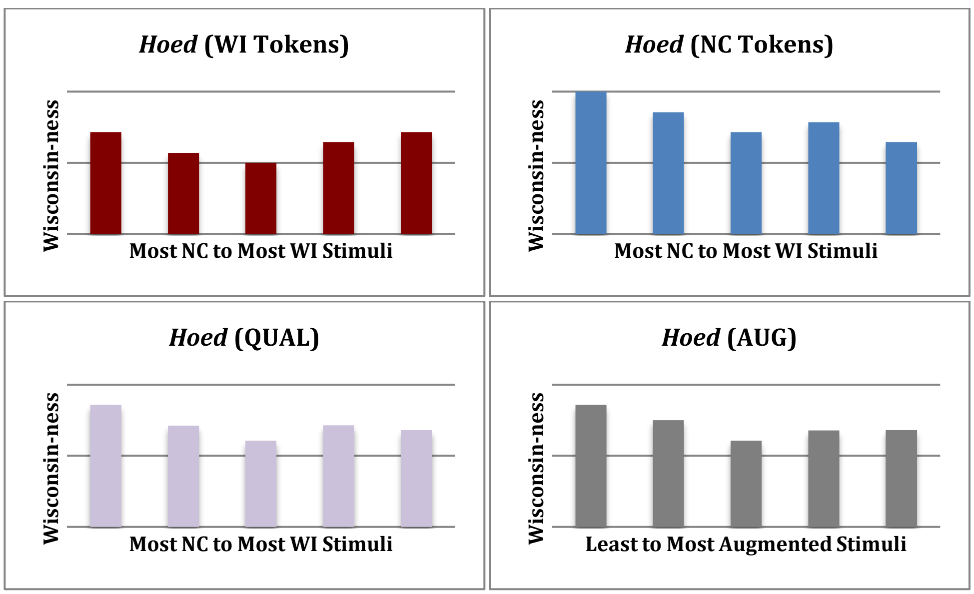

Figure 2: Averaged judgments to *hoed* according to the WI (top left) and NC (top right) tokens' continua and according to the QUAL (bottom left) and AUG (bottom right) averaging measures

The data in Figure 2 is given to illustrate how trend line information was collected. This trend line information is reported in Table 4 for all three words. In several ways, Table 4 contradicts our expectations. The slope values should be negative for WI and positive for NC, indicating that tokens with more canonically Wisconsin formant values would be judged on average as sounding more canonically Wisconsin. However, NC's slope values were negative for all words, and WI's slope was positive for *hoed*. These observations are enriched by discussing R^2 coefficients. WI's *hoed* value is extremely small, making the surprising negative slope less problematic. However, NC's *who'd* has a reasonable R^2 coefficient, and for *hoed* the value is tremendous, such that the positive slopes for these words remains an issue. For every word, there are large discrepancies between responses to WI and NC.

Table 4: Trend line information for aggregate judgments

Word	**Trend**	**WI**	**NC**	**QUAL**	**AUG**
Hoed	Slope	pos	neg	pos	pos
	R^2	0.01	0.82	0.01	0.82
Hood	Slope	neg	neg	neg	neg
	R^2	0.92	0.13	1.00	0.80
Who'd	Slope	neg	neg	neg	neg
	R^2	0.32	0.02	0.17	0.01

Next we compare QUAL to AUG in an attempt to isolate the influence of formant values as opposed to the influence of augmentation in audio quality. For *hoed*, both QUAL and AUG are negative; for *hood* and *who'd*, both QUAL and AUG are negative. This suggests that neither QUAL nor AUG can predict the slope of the trend line. Other problems emerge when considering R^2 values. The QUAL correlation is weakest for *hoed,* which is to be expected given that Wisconsin does not participate in /o/ fronting. Thus, for *hoed*, AUG explains the data much better than does QUAL. Surprisingly, the QUAL correlation for *hood* fits with perfect correlation, and yet AUG still explains a tremendous amount of variation. The values for *who'd* are more what I was expecting: a reasonable R^2 value for QUAL and a tiny one for AUG. The long and short of this analysis is that the augmentedness apparently played nearly as large a role in participant answers as formant values did.

3.3 Connection

We now turn to the connection portion of the analysis. Here, the perception data are considered dependent variables of the production data. This can be seen in Figure 3. The X-axis of each chart consists of the six participants ranked according to their average representativeness of Upper Midwestern English formant values for a given word. The Y-axis of each chart consists of the trend line R^2 correlation coefficients taken from the production data. For *who'd* and *hood*, there is rather weak correlation, much like that between formant value and judgment score reported above in the production data. For *hoed*, the pattern is stronger, though impressionistically the data show a pattern that is not linear.

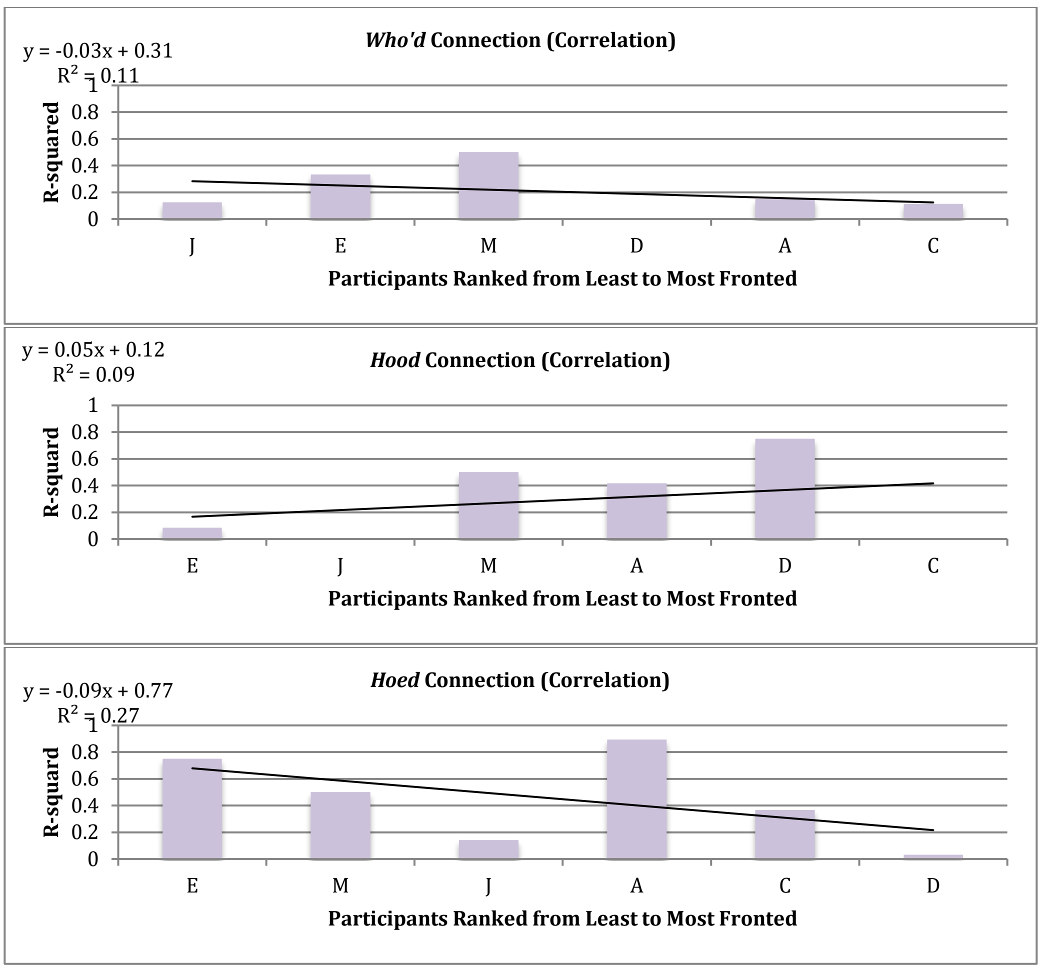

Figure 3: Connections between speaker frontedness in production and R^2 correlation coefficients from perception judgment trend lines (QUAL)

I repeated this analysis using trend line slope as opposed to trend line R^2 value. These are presented in Figure 4. Here, *hoed* shows the strongest connection while *hood* and *who'd* show the weakest. I should point out that Figure 3 and Figure 4 each only look at one aspect of the trend lines but cannot consider both at the same time. This may make Figure 3 deceiving, since participants vary in their trend line slopes.

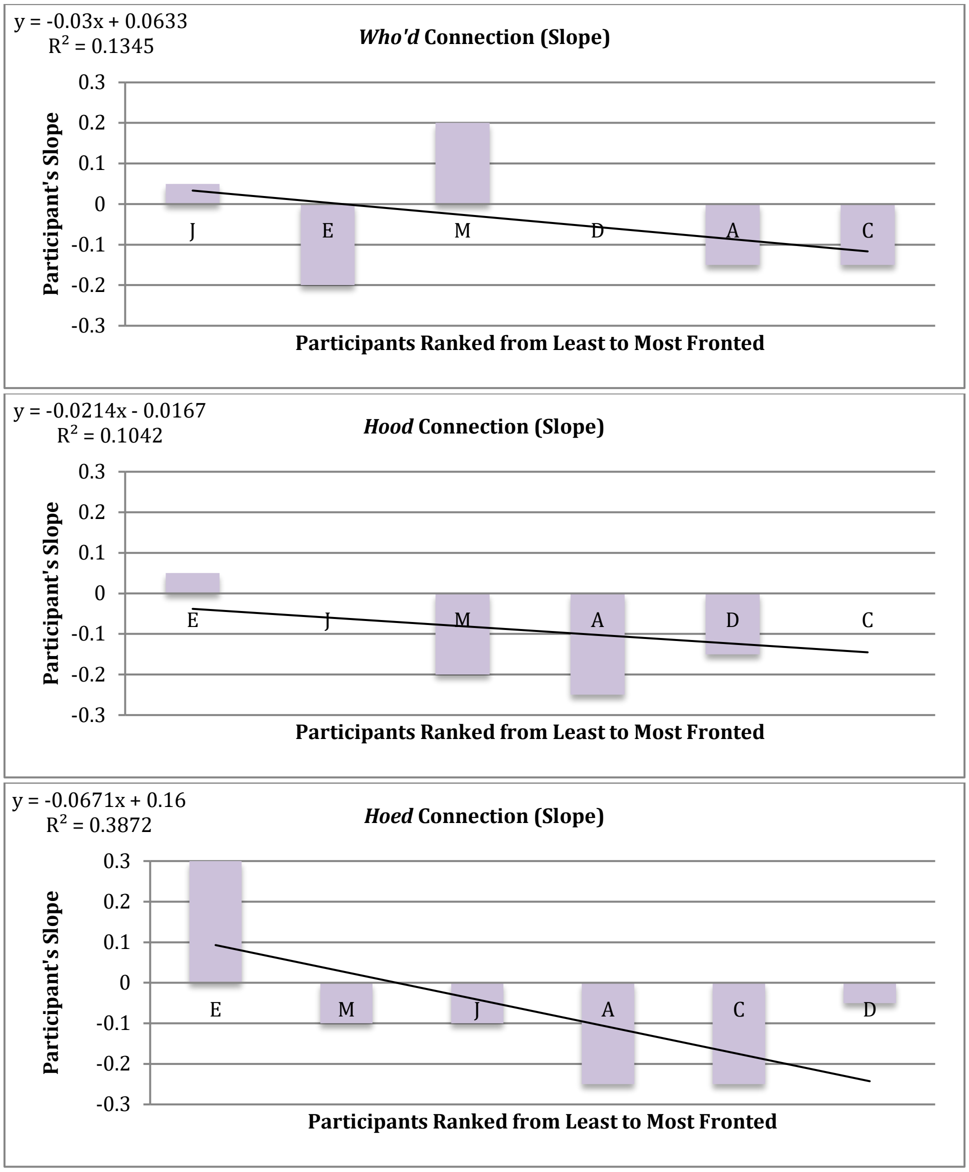

Figure 4: Connections between speaker frontedness in production and slope values from perception judgment trend lines (QUAL)

4 Discussion

The production data conform largely to expectations: my participants showed fronting of /u, ʊ/ but not /o/. One noteworthy exception is that my /ʊ/ results seem further forward than that of other researchers. The perception data I gathered is more complicated and less conclusive. There were generally weak and sometimes counterintuitive patterns for related sets of audio stimuli. Speakers were no better at discerning steps in *hoed* than in *who'd*, but they were remarkably adept at discerning steps in *hood*. But this is only on the whole, as judgments for WI and NC were not particularly similar. Figure 3 shows the interesting finding that, for *who'd* and *hood*, being more fronted meant that judgments were less predictable/consistent than average, whereas for *hoed* being more fronted meant that judgments were more predictable/consistent than average. It has been suggested to me that salience might be at play here – perhaps *hoed* is socially salient while the others are not (or vice-versa). I leave this to further research.

Connecting production with perception is not as simple as I had anticipated. The chief difficulty is that high correlation did not correspond to slope direction, making an analysis of correlation alone unreliable. There were patterns indicating a connection between frontedness and perception judgments, but they were inconsistent: some participants had high correlation and a positive slope, meaning that they judged the tokens as expected, whereas some participants had high correlation and a negative slope, meaning that they judged the tokens opposite our expectations. I believe that grouping speakers together according to inter-speaker production hallmarks might mitigate these issues, though of course this would also lose the individual basis that motivated this work in the first place.

Returning to my original hypotheses, we find that they are all in need of modification, some more so than others.

1. The group of participants perceived changes best in *hood*, then *who'd*, and then *hoed*.
2. The group of participants fronted *who'd* to some extent, but they strongly fronted *hood* and even fronted *hoed* a little.
3. Individual participants producing a more fronted back vowel did not necessarily better perceive differences in F2 for that vowel than individual participants with less fronted back vowels did – the connection was different word by word.

In this project, I was chiefly interested in connecting production and perception on an individual level. It was decidedly exploratory, and so some aspects of the methodology proved much less valuable than others. I would conduct another study as a follow-up to the present study.

The most important adaptation in my next study would be in the perception judgment. I asked participants essentially whether or not a speaker sounded like they were from Wisconsin. But "not Wisconsin" was almost certainly too broad. For one thing, Wisconsinites in one part of the state speak differently from others, and the formant values I considered as the standard for comparison were representative of Wisconsin. Additionally, the task of choosing whether someone is not from X is probably more difficult than choosing whether someone is indeed from Y. In the follow-up study, rather than giving participants a four-point scale based solely on "Wisconsin," I would present a map like that of Plichta & Preston (2005). The map I would use is below in Figure 5. This map consists of three Northern, two central, and three Southern cities.

Figure 5: Potential cities to be used for the next experiment

I would of course need to narrow my focus down in this next study. The first consideration I would make is whether using both WI and NC is really the right choice. It introduced more complication than I had anticipated. While that is not necessarily bad, I openly admit that there are theoretical ramifications that I don't yet understand. I would not know how to tease out the role of formant values versus augmented sound if I were to use one speaker, but it may simply be biting off more than I can chew to use two voices.

Whether I used two speakers for the test stimuli or just one, I would still use NC's formant values as end values for the WI continua. That is, instead of arbitrarily selecting formant modification step values and only changing one formant in successive steps, I would ensure that all three formants were changed in the amount that gave WI_4 as nearly as possible the formant values for NC_0 (and vice-versa, if using NC). This would create stimuli that sound more natural and are more authentic in vowel quality.

Another change in the next version of this research would be including every test token twice instead of just once (which would also call for a commensurate increase in the number of distractor tokens). This follows Plichta & Preston's (2005) *guide* study as well, and it would lend greater reliability to the responses collected from participants.

I would very much like to gather speaker judgments from other points on the map in Figure 5. Since Janesville is third from the top and my participants this round came from near there, getting participants from Nashville, third from the bottom, feels most appropriate. An alternative would be to gather data from central Ohio and western North Carolina, as Jacewicz, Fox & Salmons (2011) did. This may simply be out of my reach for the time being, and it would probably be more prudent to call for other researchers in those areas to follow up on my results with speakers from those areas.

Finally, I do not believe that the production data would need to come from the narrative reading – I settled on the latter here merely for the sake of recruiting more comfortably. Instead, a word-in-frame reading list might be more prudent. For example, eliciting all possible *h_d* words (*heed, hid, hayed,* etc.) in the frames "Please say _____ for me," "We asked if _____ was the right word," and "Cite _____ twice" may be sufficient.

With these considerations, I would have a deeper look at how BVF is distributed acoustically and perceptually across the North-South line in American English.

Appendix A

Billy Hudd loved gardening. Working in the dirt was the most fun way to play for him. If Billy hid this passion, or tried to, it was unsuccessful. For Christmas, his sister asked for candy, but he asked for a hoe. Alas, his parents, who'd simply hemmed and hawed (and then hawed and hemmed) paid no heed – they wouldn't even let him haggle for one. Billy just had to find another way to get the tool, but nothing was coming into his head.

There was an elderly lady in Billy's neighborhood named Gertrude Hode. Billy thought of her as an old hag, and his parents made sure he always hayed the animals she kept in the backyard. But she loved to give surprises, and she had one for Billy. She saw that he'd been playing with a rickety old builder's hod, and she admired the way Billy hoed his garden faithfully with it. She knew the thing Billy wanted, and she had one. After she hayed the goats, she decided to head as quickly as possible to the Hudd residence and make a delivery.

When Billy woke up the next day, there was a gift on the doorstep: the gardening instrument of his dreams. "Aha!" thought Billy, who'd already forgotten his run-down hod. "This is the best day of my life! Who'd have thought – she's not a hag after all. What was the likelihood of that?" Right away, Billy wrote a letter apologizing to his old neighbor and thanking his new friend. Billy hid it unsneakily on her doorstep. From that day on, Billy hayed the animals without grumbling, and he neither hemmed nor hawed about it.

"Time to put on my gloves and pull up my hood and work until I've hoed out this entire garden. Hey, who wants to join in?" Billy hid the desire to garden no longer. He'd never been so glad. And Billy Hudd never picked up the old hod again.

Appendix B

Time to read a little tale
that gives a little nod
to how playtime got a little stale
for Billy and the hod

Billy Hudd loved gardening. Nothing was more fun in his mind than playing in the dirt in his neighborhood. If Billy hid this passion, or tried to, it was unsuccessful. For Christmas, his sister asked for candy, but Billy asked for a hoe. Alas, his parents simply hemmed and hawed (and then hawed and hemmed). They paid no heed. Billy just had to find another way to get the tool, but nothing was coming into his head.

There was an elderly lady in Billy's neighborhood named Gertrude Hode. Billy thought of her as an old hag, and his parents made sure he always hayed the animals she kept in the backyard. But she loved to give surprises, and she had one for Billy. He'd been playing with a rickety old builder's hod, and she admired the way Billy hoed his garden faithfully with it. She knew the thing Billy wanted, and she had one. Who'd help the boy out if not her? After she hayed the goats, she decided to head quickly to the Hudd residence and make a delivery.

When Billy woke up the next day, there was a gift on the doorstep: the gardening instrument of his dreams. "This is the best day of my life," thought Billy, already forgetting his run-down hod. "Who'd have thought – she's not a hag after all. I'm so lucky I live in her neighborhood!" Right away, a good idea came into his head: Billy wrote a letter apologizing to his new friend for treating her like a hag. Billy hid the letter on her doorstep. From that day on, Billy hayed the animals without grumbling, and he neither hemmed nor hawed about it.

"Time to get out and work until I've hoed out this entire garden. Who'd have thought – my dreams have come true!" Billy hid the desire to garden no longer. He'd never been so glad. And Billy Hudd never picked up the old hod again.

References

Ash, Sharon. 1996. Freedom of movement: /uw/-fronting in the Midwest. In Jennifer Arnold (ed.), *Sociolinguistic variation: data, theory and analysis: Selected papers from NWAV 23 at Stanford*, 3–25. Stanford: CSLI Publications.

Baranowski, Maciej. 2008. The fronting of the back upgliding vowels in Charleston, South Carolina. *Language Variation and Change* 20(3):527–551.

Clopper, Cynthia, David Pisoni & Kenneth de Jong. 2006. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America* 118(3):1661–1676.

Evans, Betsy & Dennis R. Preston. 2001. When it's not nice to be normal: What's missing from normalized data? *University of Pennsylvania Working Papers in Linguistics* 7(3):61–65.

Fridland, Valerie & Kathy Bartlett. 2006. The social and linguistic conditioning of back vowel fronting across ethnic groups in Memphis, Tennessee. *English Language and Linguistics* 10(1):1–22.

Fridland, Valerie, Kathryn Bartlett & Roger Kreuz. 2004. Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change* 16(1):1–16.

Hagiwara, Robert. 1997. Dialect variation and formant frequency: The American English vowels revisited. *The Journal of the Acoustical Society of America* 102(1):655–658.

Harrington, Jonathan. 2012. The coarticulatory basis of diachronic high back vowel fronting. In Maria-Josep Solé & Daniel Recasens (eds.), *The initiation of sound change: Perception, production, and social factors*, 103–122. Amsterdam: John Benjamins.

Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2008. Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study. *The Journal of the Acoustical Society of America* 123(5):2825–2835.

Jacewicz, Ewa, Robert Allen Fox & Joseph Salmons. 2011. Cross-generational vowel change in American English. *Language Variation and Change* 23(1):45–86.

Kleber, Felicitas, Jonathan Harrington & Ulrich Reubold. 2011. The relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech* 55(3):383–405.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Labov, William. 2005. LSA 208: Chain shifts in progress. Course taught at the 2005 Linguistic Society of America Institute, Cambridge, MA.

Natvig, David. 2014. Perceptions of prevelar and preapical /æ/ as markers in American English dialects. Poster presented at the 19th Mid-continental Phonetics and Phonology Conference (MidPhon 19), 12–14 September 2014.

Plichta, Bartłomiej & Dennis R. Preston. 2005. The /ay/s have it. In Tore Kristiansen, Nikolas Coupland & Peter Garrett (eds.), *Acta Linguistica Hafniensia 2005* (Subjective processes in language variation and change):107–130.

Rosenfelder, Ingrid, Joe Fruehwald, Keelan Evanini & Jiahong Yuan. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite. http://fave.ling.upenn.edu (accessed November 11, 2016).

Thomas, Erik R. 1989. Vowel changes in Columbus, Ohio. *Journal of English Linguistics* 22(2):205–215.

University of Wisconsin–Madison
Department of English
7195G Helen C. White Hall
600 N Park St
Madison, WI 53706

jibson@wisc.edu