

February 8, 2006  
Linguistics 500

## Article 2 Summary

Peterson, Gordon E. and Harold L. Barney. (1952). "Control Methods Used in a Study of the Vowels." *The Journal of the Acoustical Society of America*, Vol. 24, No. 2, 175-184.

**Main research question:** What is the general relationship between the vowel phoneme intended by a speaker and that identified by a listener? How do these relate to the acoustical measurements of the formant or energy concentration positions in the speech waves?

**Specific question tested:** In a list of monosyllabic words beginning with [h], ending with [d] and differing only by the vowel in the middle, how well can listeners identify the intended vowel when spoken by a number of speakers with different dialects of American English? How do the results of the listening test compare with the physical properties of each vowel as measured by a sound spectrograph?

**Experimental procedure:** A list of words (as described above) was presented to 76 different speakers of different American English dialects, including men, women, and children, and their utterances were recorded. The word order was randomized for different speakers, and each speaker had to pronounce two different lists. For part one of the experiment, the recordings were played back to an audience of 70 adult listeners over 8 sessions (32 listeners had also been speakers). The listeners were seated randomly in an auditorium, given pads containing 200 lines of the 10 words, and asked to draw a line through the one word they heard for each playback. For part two of the experiment, the recorded lists were played to a sound spectrograph, which charted their acoustic properties. Measurements for the formant frequencies were taken for each speaker. Comparisons were then made between the results of the listening test and the spectrograph measurements.

**Results of the experiment:** For the listening test, listeners identified the ten different vowel sounds with varying degrees of consistency. Out of 152 playbacks for each word, the number of unanimous correct classifications by all 70 listeners was recorded. For [i], 143 of the playbacks were unanimously identified as [i], 138 were unanimous for [ɜː], 115 for [æ], 109 for [u], 76 for [ʌ], 76 for [v], 74 for [ɪ], 52 for [ɛ], 34 for [o], and 9 for [ɑ]. For speakers who were also listeners, if they did not differentiate clearly between the vowels while speaking them, they also had a higher degree of discrepancy from the rest of the listeners in distinguishing those vowels. It was also found that the position of the listener in the auditorium affected responses. For the

second part of the experiment, it was found that when the formant's principal components were symmetrical around one dominant component, there was less ambiguity in choosing the formant frequency than if the distribution was asymmetrical. If the fundamental frequency was already very high, this also made it difficult to define a formant frequency. In mapping the frequencies of the vowels by different speakers, there was considerable overlap between [ɜ˞] and [ɛ], [ɜ˞] and [ʊ], [ʊ] and [u], and [ɑ] and [o].

**Authors' interpretation of results:** It seems that [i], [ɜ˞], [æ], and [u] were generally well understood. When these vowels were misinterpreted by listeners, they usually answered with a vowel nearest on the vowel loop to the vowel intended by the speaker. Along the same lines, [ɑ] and [o] were often confused by listeners, as were [ɪ] and [ɛ]. These substitutions conform to present dialectical trends in American speech, where there is a shift toward smaller mouth openings and higher vowels in speech. The authors guess that [i] and [u] had high intelligibility because they are the terminal positions of the tongue in the mouth, and extreme positions for vowels. Therefore, they are harder to confuse since there are not other vowels similarly formed. The [ɜ˞] was highly intelligible because it was the only vowel with a retroflex, distinguishing it from the others. The [æ] was probably correct most of the time because listeners who would tend to use this sound interchangeably with [ɛ] would identify words they heard pronounced with either an [æ] or an [ɛ] as [æ]. In comparing the measured formant frequencies of the vowels with the answers given by speakers, the sounds with the most measured overlap with another sound were the ones that the listeners had the most trouble distinguishing. The authors' conclusion was that vowel production is not a random process; it depends on the individual's language experience.

**Problems and Questions:** This seems to be a very well-designed study. The only part that could potentially be misleading is how the experimenters used just one speaker to diagram the vowel loop. To get a more accurate and representative picture, they could have diagramed an average of at least a few different speakers. There are different dialects represented in the study, so how did they choose one "correct" dialect to diagram the loop? One other thing that could have improved the study would be to have another group that did the same procedure with different consonants on either side of the vowels. This would ensure that the consonants were not influencing the vowel production or perception. Other than that, the parts of the article dealing with equipment calibrations and measurement techniques were a bit technical for me to understand, but that is because of my lack of background in the area.