

## The effect of outliers on the perception of sound change

It is a common statistical practice to remove outliers in order to improve the validity of the mean of a sample distribution—isolated items that are several standard deviations from the mean. However, the role of outliers is a crucial unknown in the transmission and incrementation of linguistic change. Acoustic studies of sound change in progress commonly encounter outliers in the direction of the change. Language learners may do the equivalent of discarding outliers, but they may instead weight outliers more heavily than tokens in the main distribution as cues to the direction of change.

To resolve this question experimentally, the word *bad*, as originally pronounced by a young South Philadelphian woman, was modified in F1, F2, duration and inglide to produce a series of six tokens along the axis of change common to the speech community, from low front to peripheral upper mid ingliding. F1 ranged from 802 to 392, and F2 from 1729 to 2534. Subjects were asked to move six buttons associated with each sound on a sliding scale so that the distances between them correspond to the similarities and differences of sounds. Of 50 subjects, 44 arranged the six tokens in the same order as formant measurements; r-correlations with a linear progression were .99 for these 44 subjects, and equally high for all 50 subjects combined.

Subjects were then asked to help in studying how people change their speech when they move from place to place, using the phonetic scale they had just established. They were asked to register their overall impression of how the same person said *bad* after living in different four different cities. The four sets of five stimuli embodied different values of *bad* in carrier phrases like “That’s bad,” “Pretty bad,” etc. Series 1 was a symmetrical distribution in low position; Series 2 had a low outlier 1.7 s.d. from the mean; Series 3 was a symmetrical distribution in upper mid position; Series 4 had a high outlier 1.8 s.d. from the mean.

Ratings of the series with a low outlier were significantly different from the low symmetrical series at  $p < .0001$ . Ratings of the series with an upper mid outlier were significantly different from symmetrical upper mid series at  $p < .01$ . Differences for the low series were considerably greater than for the upper mid series. On a best fit superposition of the rating scales on the physical measurements, the lower mid outlier created a difference of 50 units, considerably more than the mean physical difference, while the upper mid outlier created a difference of only 10 units, less than the mean physical difference. This may be attributed to the fact that the low outlier corresponds to a correction of the vernacular form in Philadelphia, and subjective reaction tests have shown that such a correction elevates ratings on the job scale of social evaluation. In both cases, results indicate that outliers cannot be discarded in determining mean values that correspond to social perception.