

Methodological trends in acoustic phonetic analysis

Emily Grabowski

University of California, Berkeley
emily_grabowski@berkeley.edu

ABSTRACT

Acoustic analysis in phonetics presents an opportunity to quantify meaningful differences in speech sounds. Recently, however, methodological research has raised concerns about the degree of variation in acoustic methodology and its impact on generalizability of the results. While researchers often have intuitions of best practices for selecting appropriate features, there is little quantitative investigation of how acoustic features are used in empirical studies. This paper presents the results of a survey of the methodology in a sample of papers from the proceedings of this conference at two time points 20 years apart, with a focus on the target contrast, parameters used, and measurement choices. The results identify themes in acoustic features across contrast types and identifies potential areas for further development of acoustic parameters.

Keywords: methodology; acoustic analysis; phonetics

1. INTRODUCTION

A central part of designing an acoustic analysis is choosing which phonetic parameters to include. Generally, acoustic correlates are established on a per-contrast basis and researchers discuss previous work when determining appropriate contrasts to use. This results in a significant opportunity for variation in the representation of speech. In addition, the appropriate acoustic features to use have been reported to vary along dimensions such as language [1,2] and speaker [3].

Variance alone is not necessarily an issue - language is a complex social phenomenon and will always have significant inherent variation [4]. However, methodological variation may also contribute to issues in replicability, false positives, and other issues that are currently the subject of discussion in the field [5,6,7].

In addition to broader methodological studies, there have been several investigations into the quality of acoustic features. One study on the validity of acoustic features attempted to relate acoustic features to perceptual and acoustic landmarks [8], while another investigated the variance of formant measurements [9].

In addition to directly establishing the quality of a specific measure, there is value in understanding how acoustic methods are used in practice. While there is opportunity for significant methodological variation, it is not clear to what degree that it is exercised in practice. In addition, very little is quantitatively known about how they are used in practice in the broader community. One notable exception is [10], which surveys the presence of various acoustic correlates and contrasts discussed in phonetic descriptions. This survey finds that a handful of correlates (including formants and voice onset time) are used broadly and that there is an imbalance in coverage between consonants and vowels. In this paper, I describe the results of a survey designed to further explore the usage of measurement methods in acoustic phonetics.

2. METHODOLOGY

The data for this study are drawn from the Proceedings of the International Congress of Phonetic Sciences (ICPhS), a major conference that represents a broad cross-section of phonetics and adjacent fields.

The sample contains all proceedings papers for the years 1999 and 2019. The most recent proceedings available at the time of writing is 2019, and 1999 was chosen to be early enough to predate the recent wave of computational tools to aid phonetic analysis. The sample contained 640 papers in 1999 and 792 in 2019, for a total of 1432 papers. Of these, 395 (155 in 1999, 240 in 2019) were determined to be relevant based on the title and the abstract as pertaining to acoustic analysis of a consonantal, vocalic, or suprasegmental contrast.

Each relevant paper was coded for the broad phonetic category under study: sonorants, obstruents, vowels, and suprasegmentals. Studies that included multiple contrasts or contrasts that crossed these categorical boundaries were coded for both categories. Papers were also coded for the acoustic parameter(s) used to measure the contrast and how each parameter was operationalized.

The next section will present several trends in the methods used in each year for three categories: vowels, sonorants, and obstruents. Laryngeal contrasts are not reported here as they constitute a significantly different research paradigm. First, the

rate of coverage of high-level contrasts (vowels, obstruents, sonorants) was calculated as the percentage of papers in the sample for each year, in order to account for trends in the overall popularity of acoustics. All other analyses are reported on a per-contrast basis, where data are normalized to the number of papers in each contrast category that year.

3. RESULTS AND DISCUSSION

In 1991, 22% of the total papers in the proceedings fell within the sample, while in 2019 the rate increased to 32%. This suggests that there is a general increase in the popularity of acoustic measurements over the time period of the study. This could be due to several different factors, including general increased interest in quantitative measurement in phonetics and access to tools that made taking certain acoustic measurements much more straightforward.

The rate of papers covering each general phonetic category is given in Figure 1. Sonorants and obstruents see an overall downward trend between the years in the study, while vowels have an upward trend. Sonorants are also the least commonly reported, while obstruents saw approximately the same number of papers addressing them as vowels in 1999. However in 2019 there is more spread in the distribution of studies across categories, as vowels increased in representation and both sonorants and obstruents decreased.

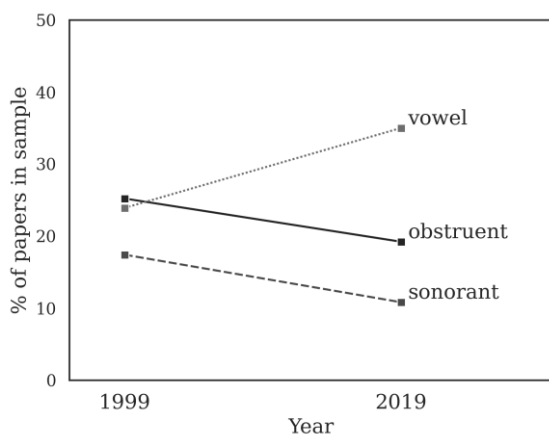


Figure 1: Percentage of papers in each sample addressing major acoustic contrasts.

While some asymmetry in representation of categories can be expected from non-methodological factors such as typological distribution of phonetic categories, those factors wouldn't be expected to influence the trends over time observed in the results of this survey. Rather, these trends over time are more likely influenced by changes in methodological practices in the community. Consonants have a wider

range of acoustic variation and the appropriate features must be established for each contrast. On the other hand, measurements for vowels, for example, are quite well understood and there are many tutorials and resources available for automatic or near-automatic analysis of relevant features. This imbalance in accessibility may be helping to encourage a trend towards more vowel analysis and less consonant analysis.

3.1. Usage of qualitative methods of analysis

The trend in phonetics has been to quantify acoustic differences where possible. Quantitative analysis helps to identify the significance and magnitude of differences between sounds and reduces the contribution of perceptual bias to results. Nonetheless, qualitative, impressionistic analysis of a segment is a valuable entry point for investigating contrasts, particularly those that do not have a clearly delimited set of descriptive acoustic features. This suggests that in this sample, a higher rate of qualitative measures for a contrasts indicates a lack of confidence in the ability of existing features to capture the dimensions of the contrast under study.

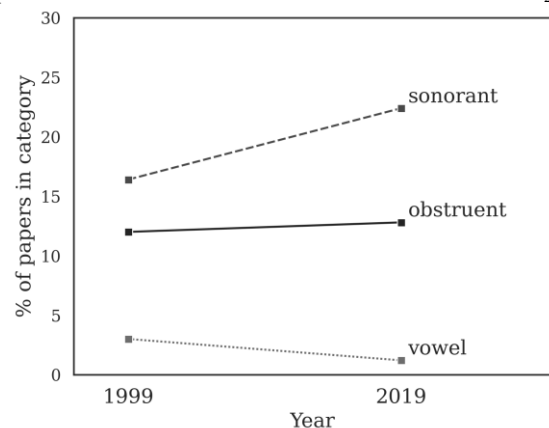


Figure 2: Percentage of papers in each category that utilize qualitative methods

The proportion of papers in each category that used qualitative measures is given in Figure 2. Vowels have the lowest rate of qualitative analyses, followed by obstruents, while sonorants have the highest rate in both years. This pattern aligns with intuitions about the relative confidence in acoustic features that represent each of these categories. Vowels and obstruents do not show much change between samples, but in sonorants there is an increase in the usage of qualitative strategies in 2019. While qualitative analysis is a valuable tool, an increase in rate of this analysis suggests that there is also an increase in uncertainty in the appropriate acoustic parameters to use to represent these segments.

The pattern reported in this section aligns with that identified in the section above: vowels are easier to

analyze quantitatively and sonorants are particularly difficult to analyze given the current set of acoustic features. These trends highlight the potential of establishing features for sonorant consonants for advancing acoustic analysis in phonetics.

3.2. Variance in feature usage and measurement

Aside from overall trends, it is also possible to investigate how the methods and measurements used within each category has changed over time. The distribution of phonetic categories across common types of acoustic features for each year is given in Figure 3, reported as the proportion of papers in each category that utilize a feature type.

Non-vowel analyses have more durational and non-acoustic approaches overall while formants are the most used feature for vowels and sonorants across both years. Formants (in particular formant transitions from the neighboring vowel) were also commonly used for obstruent analyses in 1999 but have decreased in usage. In that category, spectral moments have become the most common acoustic feature type, which appears to be a somewhat recent development.

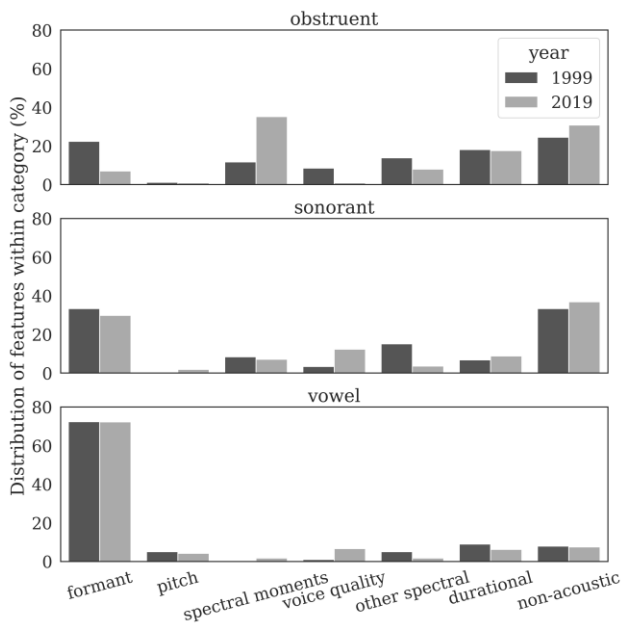


Figure 3: Distribution of sample across acoustic feature categories.

Although there are many possible acoustic parameters that have been proposed for phonetic analysis, in usage there appears to be a convergence on a few common sets of measures. This makes sense given the fact that researchers often build their analysis by reference to previous successful studies on analogous contrasts. While this helps establish a set of best practices for the field, it also makes clear

the importance of understanding the limitations of these features given that they can disproportionately impact insights from acoustic analyses. In addition, since the feature used varies by segment type, the proper analysis for contrasts that straddle these categorical boundaries is not immediately clear.

There is also potential variation in the operationalization of these measurements, in particular choices made around where to measure these features within the segment. This is addressed by Figure 4, which gives the distribution of phonetic categories across measurement types reported as proportion of papers in each sample that use each measurement type. Consonants see a move towards point and window (e.g. average value across some portion of the segment) methods, and other more complex statistical formulations, such as formulas, derivatives, etc) are less common in 2019 than in 1999. There is also an increase in trajectories (measurements of multiple time points) used in most categories, which suggests a more general trend towards dynamic measures. There is still some variation between measure types, and it should be noted that there is also significant variation at a finer level of detail within each of these categories, for example where to measure a point within a segment.

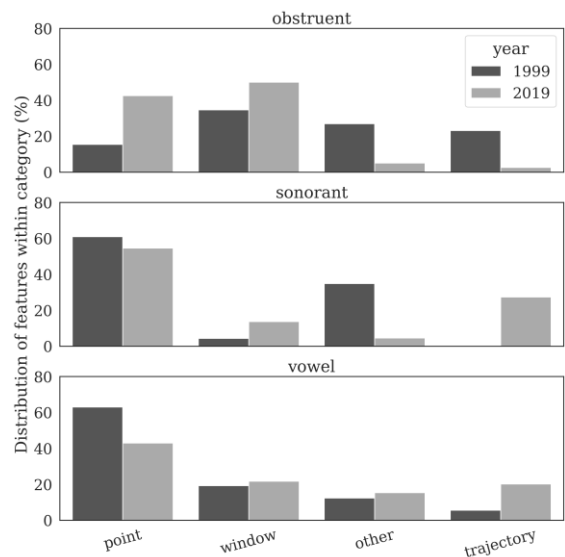


Figure 4: Distribution of sample across measurement strategies.

3.3. Methods across methods of articulation

Up until this point, the phonetic contrasts in this sample were pooled into broader categories. While this is helpful for looking at general trends, it is also informative to consider more fine-grained distinctions. From the convergence on a set of

parameters within the target phonetic categories arises the question of how to analyze data that may cross these boundaries. The properties of this distribution may give insight into the current limitations of acoustic features.

The relative frequency as well as the most used measure for each method of articulation is reported in Figure 5, pooled across years in the study. The diagonal gives the values for studies within that category and the off-diagonal values are those comparing between methods of articulation. The shading of each cell corresponds to the number of papers in that category in the paper. Where there was a tie in the most common method of analysis, both methods are noted.

There is a distinct imbalance in the data in terms of which contrasts are represented, and several potential contrasts have no or very few studies addressing them. While relative frequency may also be affected by phonetic typology, these imbalances may also be influenced by how easy or difficult it is to capture the contrast in numerical terms.

In addition, almost all studies that evaluated values across methods of articulation included either articulatory or qualitative measures in the top methods used to measure them.

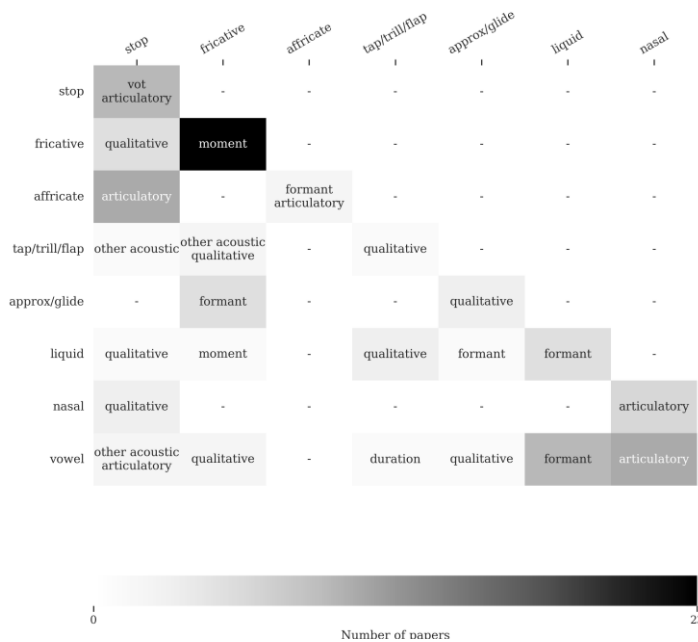


Figure 5: Most common parameters used for comparisons across methods of articulation.

The acoustic measures used most frequently are formants, which are used in both sonorant contrasts and in some obstruent analyses (by way of formant transitions). Formants are a relatively transparent, measurable, and widely-used acoustic measure, and from this analysis it is clear that they have been leveraged to apply to a broad range of phonetic

contrasts. However, the interpretation of formant differences still depends on the type of segment being measured, meaning that these measures may not always be directly comparable in different contexts.

The results presented here, including the significant use of non-acoustic measures, suggest that existing measures do not work equally well for all methods of articulation within the general phonetic categories.

4. CONCLUSION

This paper presents some results of a survey of phonetic methodologies in the field. Generally, there is an increase in acoustic studies over the time period in this study. In that time there is an increasing gap in how frequently different categories are represented in the sample, in particular less coverage of consonants and more of vowels. Consonants see more qualitative analyses than vowels, and sonorants in particular see an increase in qualitative measurement between the two study years. There is some convergence on a type of parameter, and to some extent the type of measurement, within each broad phonetic category. However, looking within categories at cross-categorical comparisons found a high rate of non-acoustic measures, suggesting that the common measures for consonants are not richly descriptive enough to capture all contrasts of interest in the study.

These results suggest there is generally less confidence shown in consonant parameters, particularly sonorants: they see a relative decrease in coverage and a higher rate of qualitative description. In addition, consonants see the most non-spectral parameters (such as durational or impressionistic measures) and more shift in parameterizations and measurement methods over the years.

While quantitative analysis has contributed to significant progress in phonetic analysis, this survey demonstrates that there are still questions about best practices in acoustic methodology in these cases, and that falling back on qualitative measures is still quite common particularly for less common contrasts. These observations indicate that there is room for further consideration of acoustic features that can be generalized across contrasts, and further understanding the dynamics of acoustic methodology in phonetics.

5. REFERENCES

- [1] T. Cho, D. H. Whalen, and G. Docherty. 2019. Voice onset time and beyond: Exploring laryngeal contrast in 19 languages, *Journal of Phonetics* 72, 52–65.
- [2] A. Jongman, S. E. Blumstein, and A. Lahiri. 1985. Acoustic properties for dental and alveolar stop consonants: a cross language study, *Journal of Phonetics*, 13(2), 235–251.

- [3] K. Amino and T. Arai. 2009. Speaker-dependent characteristics of the nasals, *Forensic Science International*, 185, 21–28.
- [4] Grieve, J. 2021. Observation, Experimentation, and Replication in Linguistics. *Linguistics* 59 (5), 1343–56.
- [5] Roettger, T. 2019. Researcher Degrees of Freedom in Phonetic Research. *Laboratory Phonology* 10 (1).
- [6] Sönning, L., Werner, V. 2021. The Replication Crisis, Scientific Revolutions, and Linguistics. *Linguistics* 59, no 5, 1179–1206.
- [7] Winter, B., Grice, M. 2021. Independence and Generalizability in Linguistics. *Linguistics* 59 (5), 1251–77.
- [8] Gradoville, M.S. 2011. Validity in measurements of fricative voicing: evidence from Argentine Spanish. *Proc. 5th Conference on Laboratory Approaches to Romance Phonology*, 59-74.
- [9] Kendall, T and C. Vaughn. 2015. Measurement variability in vowel formant estimation: a simulation experiment. *Proc 18th ICPHS Glasgow*.
- [10] Whalen, D. H., Dicanio, C., and R. Dockum. 2022. Phonetic Documentation in Three Collections: Topics and Evolution. *Journal of IPA* 52, no. 1, 95–121.