Phonetic aspects of many languages have been documented, though the breadth and focus of such documentation varies substantially. In this survey, phonetic aspects (here called ‘categories’) that are typically reported were assessed in three English-language collections – the Illustrations of the IPA from the Journal of the International Phonetic Association, articles from the Journal of Phonetics, and papers from the Ladefoged/Maddieson Sounds of the World’s Languages (SOWL) documentation project. Categories were defined for consonants (e.g. Voice Onset Time (VOT) and frication spectrum; 10 in total), vowels (e.g. formants and duration; 7 in total) and suprasegmentals (e.g. stress and distinctive vowel length, 6 in total). The Illustrations, due to their brevity, had, on average, limited coverage of the selected categories (12% of the 23 categories). Journal of Phonetics articles were typically theoretically motivated, but 64 had sufficient measurements to count as phonetic documentation; these also covered 12% of the categories. The SOWL studies, designed to cover as much of the phonetic structure as feasible in an article-length treatment, achieved 41% coverage on average. Four book-length studies were also examined, with an average of 49% coverage. Phonetic properties of many language families have been studied, though Indo-European is still disproportionately represented. Physiological measures were excluded as being less common, and perceptual measures were excluded as being typically more theoretical. This preliminary study indicates that certain acoustic properties of languages are typically measured and may be considered as an impetus for later, fuller coverage, but broader consensus on the categories is needed. Current and future documentation efforts would benefit these considerations being addressed.

1 Introduction

Language documentation has received increasing attention within linguistics in recent years. A major impetus is the rapid decline in the number of languages being spoken (e.g. Hale et al. 1992, Nettle & Romaine 2000); indeed, the number of language families that are threatened
is quite large as well (Whalen & Simons 2012). Phonetic documentation is a part of that effort, but it is one that has not received widespread attention. The phonetic structure of a language, as with any other linguistic level, is dauntingly complex, but documentation efforts are still worthwhile. Flemming, Ladefoged & Thomason (2008: 465) expressed their aim as ‘to demonstrate how current techniques of phonetic investigation can be applied so as to be able to construct a basic, but reasonably comprehensive, set of materials for an archive of the sounds of an American Indian language’. They further add that ‘the failure to be able to provide everything does not mean that we should not try to provide as much as we can’ (ibid.). The present paper presents a preliminary survey of what has been done in a portion of the English-language peer-reviewed literature in terms of phonetic documentation.

The details of what constitutes ideal, or even adequate, documentation has also received attention (e.g. Himmelmann 1998, Woodbury 2003, Austin 2014). To a certain extent, conceptions of adequacy of documentation change as technology changes. For millennia, a written transcription was all that was available, leading to debate about whether language change was based on letters or sounds (e.g. Robins 1968: 186). With the advent of recorded sound, far more detailed documentation became possible, especially for phonetic aspects of language. Many recordings were made of minority languages on Edison cylinders (Bauman 2011, de Graaf 2013), and some of those have been restored for current use (Haber 2014). As recording technology improved and storage capacity increased, expectations for the amount of material also increased. Although no language will ever be fully documented, the boost in capacity has certainly made it possible to study phonetics in greater detail.

The articulatory aspects of speech are equally important in understanding phonetics, and the history of the shift due to technology is similar to that for acoustics. Introspection has provided substantial information about the physical basis for the sounds that are perceived, even from our earliest linguistic records (Kemp 1995, Kiparsky 1995). As with sound, devices that make a more permanent record and allow for more quantitative analysis have enhanced our ability to gain deeper insight. Early studies included not only examinations of majority languages (e.g. Rousselot 1897–1908, Viëtor 1898) but field studies as well (Goddard 1905). Although numerous physiological studies are performed currently (including those outside the laboratory; Whalen & McDonough 2015), articulatory data are not yet a standard component of language documentation.

Indeed, the specifics of phonetic documentation in general have received only sparse attention (e.g. Ladefoged & Maddieson 1996a, Ladefoged 2003, Bhaskararao 2004), and there has been no comprehensive list of phonetic features that should be expected in a documentation project. Prior to establishing such a list, it is advisable to determine which features have been included in existing studies. To that end, we examined three relevant collections: Articles in the Journal of Phonetics (JPhon), the Illustrations of the IPA in the Journal of the International Phonetic Association (JIPA), and the various ‘phonetic structures’ articles written by Peter Ladefoged, Ian Maddieson, and their colleagues in the Sounds of the World’s Languages project (SOWL; funded by the US National Science Foundation). The first two sources are not exclusively concerned with quantitative phonetic documentation, of course, but nonetheless there is substantial material there. In addition, there are many articles in JIPA that are not part of the Illustrations of the IPA series but which qualify as phonetic descriptions. We restricted our attention to the Illustrations of the IPA because they are attempts at describing all of the relevant phones of a language or dialect, at a minimum via transcription. However, acoustic measurements can be included, and the coverage can be seen as an indication of what measurements are typical and/or expected. The SOWL collection, which appeared in a number of venues, provides a comparison in that these articles were primarily intended to provide greater descriptive phonetic coverage; being journal articles, however, they were necessarily restricted in scope, and the scope that was selected can be compared with the other collections.

It might seem inappropriate to examine two collections that are not devoted to the topic at hand – phonetic documentation – but there are no journals that specialize in that domain.
Each of the journals chosen does, as we will see, include phonetic documentation in some of their articles, even if such documentation is not the main purpose for either journal. For the *Journal of Phonetics*, theoretical issues are more central to the articles that are published. With the Illustrations of the IPA, the main goal is to show the segmental (and possibly suprasegmental) inventory of a language or dialect. Completeness in the description of phonetic attributes is neither expected nor encouraged. The instructions for the Illustrations of the IPA at one time stated: ‘In general, a submission to this section of *JIPA* should be relatively brief and not a fully-fledged article on the phonetics of the language’. In response to changing expectations for phonetic research in the field at large, this has now been changed to: ‘Although submissions to this section of *JIPA* can be relatively brief, fully-fledged articles on the sound system of the language described, providing additional detail (including supporting evidence from acoustics, articulation or perception), are strongly encouraged’ (https://www.cambridge.org/core/services/aop-file-manager/file/575ac14dccb7d317902c964b7f/IA-ifc.pdf, accessed on 15 June 2020). Illustrations are expected to provide a translation of the story ‘The North Wind and the Sun’ or another short representative text as spoken by one native speaker of the language described. Despite the historical focus on the IPA, quantitative analyses have appeared in an increasing number of Illustrations over the years. One motivation for the quantitative analyses is to provide some evidence demonstrating the accuracy of the phonetic transcription chosen for the language.

Finally, we will compare documentation coverage from these three collections to a few book-length phonetic studies in English. These do not constitute a collection, being separately published, but the expectation is that they would have greater coverage and thus be a point of comparison to the collections.

### 1.1 Collections

Our three main collections were limited to publications in English, in order to make the survey feasible with our resources and appropriate in size to the exploratory nature of the study. There are examples of collections in other languages, such as publications in Spanish covering indigenous languages of Latin America (e.g. http://repositorio.ciesas.edu.mx and http://repositorio.pucp.edu.pe/index/) and in Chinese covering various languages (http://navi.cnki.net/knavi/JournalDetail?pcode=CJFD&pykm=ZYUB). Dissertations and theses (in English as well as other languages) often include documentation, but they were beyond the scope of this study as well. Future work on phonetic ontologies should include assessment of these additional resources as well.

Our first collection was the Illustrations of the IPA. The first of these appeared in 1989 (though ‘specimens’ had appeared in *Le Maître Phonétique*; see Hirst 2010). We have taken these articles as our defining collection because they are intended to provide a sketch of the sound system of the language. Being the outlet for the IPA itself, the transcription is the primary objective. Nonetheless, acoustic measurements have accompanied many of the Illustrations, and we take the types of measurements made and more general classes of phonetic aspects of the languages as our first pass at a set of ‘categories’ that one might expect in the phonetic documentation of a language. Categories were largely types of acoustic measurements, but some of them were more inclusive (see next section, ‘Method’, for more details).

The second collection was selected from the *Journal of Phonetics*. Although this journal was first published in 1973, the earliest of that journal’s articles that we included appeared in 1984. This journal has a largely theoretical approach to phonetics, and thus documentation is usually incidental (though see SOWL examples in the next paragraph). We limited our analysis to the categories derived from the Illustrations of the IPA collection. This meant that some of the features that received a fair amount of attention in the *Journal of Phonetics*, such as voice quality, sex differences, and coarticulation, do not fit neatly into the categories we chose. For each of those categories, it can be argued that they remain on the ‘theoretical
research’ end of the spectrum rather than the ‘documentation’ end. The relevant measures for voice quality are still debated, and very few attempts have been made to provide a single acoustic metric for coarticulation. Until those important aspects of speech yield to more systematic and widely used measures, we would argue that nondistinctive voice quality measures should continue to be counted as theoretical research rather than documentation. Distinctive voice quality, at a minimum, should ultimately be included, as it was in many of the SOWL publications.

The third collection was the Sounds of the World’s Languages. The first SOWL phonetic structures article appeared in 1993. There were 25 published studies in all, most of which appeared in peer-reviewed journals (six appeared only in the UCLA Working Papers in Phonetics). Two of them appeared in the Journal of Phonetics and are counted twice in our statistics. As expected, these articles provided more coverage of our phonetic categories than the other two collections because it was their intent to provide such coverage.

2 Method

We surveyed the Illustrations of the IPA to extract a set of emergent categories. These are listed in Table 1. Categories were selected, with the aim to be a starting point for further discussion, in two stages. First, we discussed what kinds of features would be reasonable for phonetic sketches to include (VOT, vowel formant measures, etc.). Following this, we analyzed a preliminary sampling of selected sketches from the Illustrations series, working on the assumption that if some sketch included a particular phonetic measure, it would be a reasonable starting place to ask if others also included it. The final categories had to occur in more than one article, but there was no attempt to make them conform to a hierarchy. While many of the categories are self-evident, specific ones emerged in the analysis of the different articles. For instance, ‘interactions’ for either Vowels or Suprasegmentals indicated that the author(s) examined how two of the self-evident categories interacted, e.g. the relationship between vowel duration and dispersion or the relationship between tone and intonation. The category ‘vowel features’ was used as a general grouping for analyses of vowel features like phonation, ATR, and nasalization. Thus some of the categories overlap, and some are open to interpretation. We will discuss limitations and future directions in Section 4, ‘Discussion’, below.

<table>
<thead>
<tr>
<th>Table 1 Phonetic categories surveyed.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consonants</strong></td>
</tr>
<tr>
<td>Voice Onset Time (VOT)</td>
</tr>
<tr>
<td>Closure duration</td>
</tr>
<tr>
<td>Voicing/Voicing duration</td>
</tr>
<tr>
<td>Formant transitions (place)</td>
</tr>
<tr>
<td>Fricative spectrum</td>
</tr>
<tr>
<td>Fricative duration</td>
</tr>
<tr>
<td>Burst characteristics</td>
</tr>
<tr>
<td>Preaspiration</td>
</tr>
<tr>
<td>Sonorants</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

We assigned a value of 0, 0.5 or 1 to each category in each paper. The extreme values represent no (0) or good (1) representation of the category in the acoustic measurements. If only partial results were presented, such as measuring formants for only a few vowels
rather than the whole inventory, or discussing the VOT of different consonant types without presenting any actual measurements, a value of 0.5 was assigned.

The interpretation of the number values assigned to these categories varied slightly for each of the collections. If an article within the *JIPA* collection included either a spectrogram or some brief measurements for a particular category, e.g. contrastive vowel length, the category was assigned a value of 1. However, owing to the greater analytical depth in the *JPhon* and SOWL collections, the inclusion of a spectrogram for a particular category would garner a value of 0.5. A value of 1 for a category in these collections necessitated the inclusion of a table of observed measurements or a figure illustrating a set of averaged values. We believe that each of the collections should be evaluated on their own terms for present purposes, although a unified approach is desirable for the future.

As indicated earlier, this set of categories was derived from the Illustrations, not from first principles. They are not a completely coherent set. The category ‘sonorants’, for instance, could include formant, duration and/or intensity measurements, but in fact, sonorant consonants were seldom measured at all. Our intent was to see how often the most common categories were used, and allow them to determine a rough estimate of coverage. If we included categories for every possible combination of duration, frequency and intensity, there would be many ‘0’ cells for every description we examined. That might be useful at some future time, but it was deemed counterproductive for this initial survey.

Some categories with zeroes should not count against a study, of course. If the target language does not have preaspiration or tone, one would not expect them to be measured. Neither is it appropriate to give credit for them, say, by giving a 1 if the language lacked the feature. In any event, we did not have the resources to track down which gaps occurred in the languages studied, so a lack of documentation of a non-existent category is currently counted as a lack of documentation, and future studies should address how to deal with such gaps. For instance, is phonetic documentation more complete if one examines only formant values for a language with a common, average size vowel inventory (e.g. /i e a o u/) than for a language with additional vocalic features, e.g. /i iː eː aː oː uː/?

Each collection was surveyed by one of the authors. Because the categories are provisional, only one judge was used. If the categories were agreed upon by the community at large, multiple judges would have been used to allow assessment of agreement. Indeed, the two articles that were included in two collections (*JPhon* and SOWL) revealed differences even there. Disagreements about coverage can be expected to remain even after better criteria are developed, and methods for accounting for discrepancies have been proposed (e.g. Banerjee et al. 1999, Kottner et al. 2011). Because of the preliminary nature of the present categories, a systematic assessment of degrees of disagreement was deemed premature. If a more definitive set of categories emerges, validation across raters would clearly be called for.

### 3 Results

#### 3.1 *Journal of the International Phonetic Association (JIPA)*

The *Journal of the International Phonetic Association* began publication in 1971 as a continuing publication of the International Phonetic Association. Two previous journals, *The Phonetic Teacher* (which began in 1886) and *Le Maître Phonétique* (1889–1970, with some years suspended), were other organs for the association. In 1989, a series of articles was begun under the rubric ‘Illustrations of the IPA’, with Peter Ladefoged (anonymously) commencing with American English, embedded within the article known as the Kiel report (International Phonetic Association 1989). This report, substantially updated the IPA, and so the intent of the new series was that the ‘IPA should be illustrated by transcriptions in a range of languages. For each language there should be a word list, with English glosses, illustrating all the major surface phonetic contrasts that occur in the language, and a connected text’
The transcription was meant to represent what was recorded, and not an idealization. It was noted at the time that ‘[a] recording of all this material should be available’ (ibid.), but ‘available’ was left unspecified. Many of the Illustrations have accompanying online recordings (https://richardbeare.github.io/marijatabain/ipa_illustrations_all.html). The story ‘The North Wind and the Sun’ was suggested as a means to save space, since no translation would need to be included. Further, each Illustration was to include notes on allophony not obvious from the transcription, lexically relevant suprasegmental detail (e.g. tone or unpredictable stress), the example passage in the original orthography (revealing a subtle bias toward languages with written traditions), and an audio recording (International Phonetic Association 1989: 77–78). Because the IPA is a transcription tool, it is unsurprising that most of the discussion has been about how to transcribe a language. Nonetheless, acoustic aspects have been and continue to be reported.

As of the end of 2018, the series has published 161 sketches, with a lull in the mid-to-late 1990s, and a burst in the past five years (see Figure 1). To examine trends in this corpus, we divided the corpus into three decade-long periods: 1989–1998, 1999–2008, and 2009–2018.

The original sample sketch of American English was a little under three pages long, and Illustrations of that length were common over the next several years. Since then, they have continually trended upward in both length and phonetic detail. For example, there were no sketches over seven pages in length in the first decade of the series, and the average page length in each period is 4.1, 7.2, and 11 pages, respectively. The longest Illustration in our survey’s time span is on Erusu (Chirkova et al. 2015), at 25 pages. Figure 2 shows a scatterplot of Illustration length by publication year.

Coverage was calculated as the percentage of articles studied that reported data for each category; articles that received a rating of 0.5 (see Section 2 above) were thus counted as half a reference. Values for the various categories ranged from 0% to 56.5%, with a mean of 11.6% across all 23 with a standard deviation of 9.3%; see Table 2 for a breakdown by category. VOT (11.8%) and voicing (15.9%) are the most commonly reported consonantal categories; formants (33.7%) and dispersion (26.0%) for the vowels; and stress (25.7%) and tone/pitch accent (24.1%) for the suprasegmentals. There is a slight tendency for coverage to increase across the three decades (1990s, 8.9%; 2000s, 10.7%; 2010s, 12.7%), but there is a
Table 2  Coverage by category for JIPA.

<table>
<thead>
<tr>
<th>Consonant categories</th>
<th>Coverage</th>
<th>Vowel categories</th>
<th>Coverage</th>
<th>Suprasegmental categories</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOT</td>
<td>12%</td>
<td>Formants</td>
<td>34%</td>
<td>Stress</td>
<td>25%</td>
</tr>
<tr>
<td>Closure duration</td>
<td>10%</td>
<td>Dispersion</td>
<td>25%</td>
<td>Length</td>
<td>7%</td>
</tr>
<tr>
<td>Voicing</td>
<td>15%</td>
<td>Additional features</td>
<td>17%</td>
<td>Tone or pitch accent</td>
<td>29%</td>
</tr>
<tr>
<td>Formant transitions</td>
<td>2%</td>
<td>Duration</td>
<td>26%</td>
<td>Intonation</td>
<td>16%</td>
</tr>
<tr>
<td>Fricative spectra</td>
<td>2%</td>
<td>Intensity</td>
<td>1%</td>
<td>Interactions</td>
<td>16%</td>
</tr>
<tr>
<td>Fricative duration</td>
<td>3%</td>
<td>Interactions</td>
<td>18%</td>
<td>Other</td>
<td>0%</td>
</tr>
<tr>
<td>Burst</td>
<td>8%</td>
<td>Other</td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preaspiration</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonorants</td>
<td>18%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2  Plot showing the increase in page length of Illustrations in JIPA in recent years.

great deal of variability here. Many Illustrations continue the tradition of containing few or no acoustic measurements at all.

The representation by language family is heavily skewed toward Indo-European (see Figure 3), accounting for 40% of all Illustrations published. This includes seven dialects of English, which is not unexpected for a research community for whom that is the most common native language. We had expected that this Indo-European overrepresentation would decrease over time, but, as can be seen in Figure 3, the proportions have hardly changed over the decades. Illustrations are not commissioned, and they do tend to reflect the proportion of phonetic research in the world’s language families.

3.2  Journal of Phonetics (JPhon)

The Journal of Phonetics has been published since 1973, generally with four issues per year. It has always had a theoretical and experimental emphasis. One of the 12 topics that are currently listed on the journal’s website as example research areas is ‘[d]escriptive phonetics pertaining to individual languages’. Phonetic documentation is thus not excluded from the journal, but it is also not prominent. Nonetheless, it is one of the major journals in the field and constitutes one of our three samples of practice in phonetic documentation.
Of the approximately 1,560 articles published, 110 were deemed to be potentially relevant to our study on the basis of the title and abstract. These were examined in more detail by two of the authors for a judgment of whether the article could count as documentation, and initial disagreements were discussed until a decision was made. Measurements of at least an entire category in the phonemic inventory were necessary at a minimum, e.g. the vowels or nasal consonants of the language. The amount of material and coverage needed was not fixed, and the decisions necessarily rested on multiple considerations. The number of speakers and tokens measured was one feature, with larger numbers, naturally, making it more likely that a ‘yes’ judgment would be made. Measurements of a single aspect of a distinction (such as perturbations of fundamental frequency (f0) by stop voicing) were generally insufficient, while more global (and common) measures such as voice onset time (VOT) would be sufficient. It is always possible to examine the interactions between different phonetic distinctions, but if such interactions were the primary goal of the article, it was excluded from consideration. For instance, a description would ideally contain measurements for all members of a particular phonological category, e.g. all stops, all nasals, or all coronal consonants. Those articles specializing on a small subset of the phonetic contrasts in a series were excluded. The presence or absence of a theoretical conclusion was not taken into account; most of the articles did indeed make a theoretical point. Articles that dealt exclusively with perception were excluded; see the ‘Discussion’ section for further considerations about perception.

Coverage for the categories ranges from 2.2% to 39.1%, with a mean of 12.5% and a standard deviation of 7.4%; see Table 3 for a breakdown by category. VOT (29.7%) and both closure duration and voicing (each 20.3%) were the most commonly reported consonantal categories; ‘vowel features’ (25.0%) and formants (22.7%) for the vowels; and tone/pitch accent (17.2%) and interactions (12.5%) for the suprasegmentals. There was an uptick in coverage between the 1990s (7.4%) and the first decade of the 2000s (16.1%), but then a slight decline in the 2010s (13.1%). As with JIPH, there is a great deal of variability in each of those ranges.
3.3 Ladefoged, Maddieson and colleagues, Sounds of the World’s Languages (SOWL)

Largely funded by the US National Science Foundation, Peter Ladefoged, Ian Maddieson and colleagues spent decades recording and describing the phonetics of as many languages as they could. The studies involved students and senior colleagues, ultimately including about 40 co-authors. Most of the resulting studies appear in the UCLA Working Papers in Phonetics (WPP), where they continue to be archived and freely accessible. Of the 25 phonetic structures articles in the UCLA WPP, 19 were published in journals as well, and they were a major source for Ladefoged & Maddieson (1996b).

As might be expected, these studies examined a substantially higher proportion of phonetic categories than those in the other two collections. On average, they covered 40.9% of our categories. They ranged from 18.2% to 68.2%, with a standard deviation of 12.9%. The most common consonantal categories were again VOT (82.0%) and closure duration (52.0%). For vowels, formants (100%) and dispersion (92.0%) were most common. For the suprasegmentals, interactions (52.0%) and length (34.0%) predominated. For two of the languages missing VOT as a category (Amis (ISO 639-3: ami) and Sele (ISO 639-3: snw)), the languages lack voicing distinctions, so, measuring VOT might seem unnecessary; indeed, descriptions of Australian languages, which typically lack a voicing distinction, do not report extensive VOT measures (e.g., Butcher 1996, Bowern, McDonough & Kelliher 2012). However, VOT measures in several languages without a voicing distinction reveal an equal number of realization patterns (Kakadelis & Whalen 2018). Table 4 shows the rates of coverage for each category.

Table 3  Coverage by category for JPhon.

<table>
<thead>
<tr>
<th>Consonant categories</th>
<th>Coverage</th>
<th>Vowel categories</th>
<th>Coverage</th>
<th>Suprasegmental categories</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOT</td>
<td>30%</td>
<td>Formants</td>
<td>23%</td>
<td>Stress</td>
<td>10%</td>
</tr>
<tr>
<td>Closure duration</td>
<td>20%</td>
<td>Dispersion</td>
<td>6%</td>
<td>Length</td>
<td>5%</td>
</tr>
<tr>
<td>Voicing</td>
<td>20%</td>
<td>Additional features</td>
<td>25%</td>
<td>Tone or pitch accent</td>
<td>17%</td>
</tr>
<tr>
<td>Formant transitions</td>
<td>5%</td>
<td>Duration</td>
<td>23%</td>
<td>Intonation</td>
<td>8%</td>
</tr>
<tr>
<td>Fricative spectra</td>
<td>5%</td>
<td>Intensity</td>
<td>6%</td>
<td>Interactions</td>
<td>13%</td>
</tr>
<tr>
<td>Fricative duration</td>
<td>3%</td>
<td>Interactions</td>
<td>5%</td>
<td>Other</td>
<td>6%</td>
</tr>
<tr>
<td>Burst</td>
<td>9%</td>
<td>Other</td>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preaspiration</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonorants</td>
<td>7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>27%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4  Coverage by category for SOWL.

<table>
<thead>
<tr>
<th>Consonant categories</th>
<th>Coverage</th>
<th>Vowel categories</th>
<th>Coverage</th>
<th>Suprasegmental categories</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOT</td>
<td>82%</td>
<td>Formants</td>
<td>100%</td>
<td>Stress</td>
<td>20%</td>
</tr>
<tr>
<td>Closure duration</td>
<td>52%</td>
<td>Dispersion</td>
<td>92%</td>
<td>Length</td>
<td>34%</td>
</tr>
<tr>
<td>Voicing</td>
<td>10%</td>
<td>Additional features</td>
<td>69%</td>
<td>Tone or pitch accent</td>
<td>30%</td>
</tr>
<tr>
<td>Formant transitions</td>
<td>34%</td>
<td>Duration</td>
<td>28%</td>
<td>Intonation</td>
<td>10%</td>
</tr>
<tr>
<td>Fricative spectra</td>
<td>28%</td>
<td>Intensity</td>
<td>2%</td>
<td>Interactions</td>
<td>52%</td>
</tr>
<tr>
<td>Fricative duration</td>
<td>20%</td>
<td>Interactions</td>
<td>78%</td>
<td>Other</td>
<td>24%</td>
</tr>
<tr>
<td>Burst</td>
<td>14%</td>
<td>Other</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preaspiration</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonorants</td>
<td>34%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Across the board, phonetic investigations into suprasegmental categories were less represented than work on segmental categories. Within the consonantal categories, VOT and closure duration were analyzed in many studies, but burst acoustics and voicing (here, degree of voicing during closure) were absent in most studies. Within the vocalic categories, most studies included an analysis of formants and vowel dispersion, but fewer studies included data on vowel duration or intensity.

Even though the sketches were designed to be as comprehensive as feasible in an article-length format, some of the studies did in fact focus on a particular aspect of the language, such as vowels (Maddieson & Gordon 1996) or clicks (Sands, Maddieson & Ladefoged 1996). Overall, the coverage of our categories in these articles is fairly good, considering that they are similar in magnitude to those of book-length studies (see next section).

The SOWL program addressed endangered languages, covering an impressive 20 families in the 25 papers reviewed. In addition, the only Indo-European language was Scottish Gaelic, a language that has received scant attention in the phonetics literature. The typological coverage within this collection was much greater than that which was observed in the other collections. We return to this point in the ‘Discussion’ section below.

### 3.4 Books

There is no ‘collection’ of books of phonetic description, but our expectation was that book-length phonetic studies, or books that contain substantial amounts of acoustic phonetic description, would have a larger degree of coverage than the articles in the three collections. We were only able to find four phonetic book-length studies written in English: for Navajo (ISO 639-3 nav; McDonough 2003), Witsuwit’en (ISO 639-3 bcr; Hargus 2007), Shipibo (ISO 639-3 shp; Elías-Ulloa 2010), and English (ISO 639-3 eng; Olive, Greenwood & Coleman 1993). The Witsuwit’en study is not exclusively about phonetics, but its more than 200 pages that are devoted to phonetics represent a greater length of description than any of the articles we have considered.

The coverage for the four books, calculated the same way as for the articles, was 56.5% for Navajo, 45.7% for Witsuwit’en, 67.4% for Shipibo, and 26.1% for English; see Table 5 for a breakdown by category. The English number is quite low because most of the covered categories received a 0.5 rating due to the lack of tables giving means and standard deviations. The authors did not intend this volume as a record in that sense; their aim seemed to be more to present a set of examples of the kinds of phonetic effects that exist in English and how they can be measured acoustically.

#### Table 5 Coverage by category for books.

<table>
<thead>
<tr>
<th>Consonant categories</th>
<th>Coverage</th>
<th>Vowel categories</th>
<th>Coverage</th>
<th>Suprasegmental categories</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>VOT</td>
<td>88%</td>
<td>Formants</td>
<td>100%</td>
<td>Stress</td>
<td>50%</td>
</tr>
<tr>
<td>Closure duration</td>
<td>88%</td>
<td>Dispersion</td>
<td>50%</td>
<td>Length</td>
<td>25%</td>
</tr>
<tr>
<td>Voicing</td>
<td>38%</td>
<td>Additional features</td>
<td>63%</td>
<td>Tone or pitch accent</td>
<td>50%</td>
</tr>
<tr>
<td>Formant transitions</td>
<td>38%</td>
<td>Duration</td>
<td>75%</td>
<td>Intonation</td>
<td>0%</td>
</tr>
<tr>
<td>Fricative spectra</td>
<td>63%</td>
<td>Intensity</td>
<td>0%</td>
<td>Interactions</td>
<td>0%</td>
</tr>
<tr>
<td>Fricative duration</td>
<td>50%</td>
<td>Interactions</td>
<td>63%</td>
<td>Other</td>
<td>25%</td>
</tr>
<tr>
<td>Burst</td>
<td>38%</td>
<td>Other</td>
<td>63%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preaspiration</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sonorants</td>
<td>75%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>88%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Coverage for suprasegmentals is lower than for segments. McDonough (2003), for example, explicitly excluded prosody from her analysis of Navajo. If we exclude our
suprasegmental category from our calculations, the percentages rise to 81.3% for Navajo, 84.4% for Shipibo, and 37.5% for English, but they fall to 40.6% for Witsuwit’en, which had extensive coverage of the effects of stress.

It is not surprising that the coverage in books was greater than that in articles, on average, but the difference is not as large as one might expect. The existence of only one book for English is, we would say, surprising. Because it is based on one (or sometimes two) speaker(s), its limitations are greater than the 26.1% coverage would suggest. If, at some future time, the numbers are weighted by the number of speakers measured, the result would be a further reduction in the coverage. Again, however, coverage of the type quantified here did not seem to be the aim of that book.

It is worth noting that documentation/description is often divided across individual, peer-reviewed articles. Thus, it is easy to find several studies of English vowels, others on English stops and still others on English fricatives. Taken together, one can make a general statement about the phonetics of English; however, none of those articles is individually comprehensive. Nonetheless, there is little effort towards assembling disparate findings in a cohesive book. Such a book would be well-cited as a reference across fields (Communication Disorders, Linguistics, TESOL), so it probably could be published, but it is hard to imagine the project of assembling these different findings as being well-funded. Further, for a well-described language, the expectations for the number of speakers and tokens would be quite high, making a systematic monograph a major undertaking. Even making use of existing acoustic corpora would require substantial effort. It seems that it has been too sizable a task to elicit such a book.

4 Discussion

The types of acoustic measurements reported in the literature have grown in number over the years, but this growth is rather organic, in that there is no published source for what one might consider a comprehensive phonetic description. As such, measurements that are both useful and readily determined from acoustics alone, such as VOT and vowel formants, are well-represented. Aspects that are less accessible, such as formant transitions, or extremely complicated, such as intonation, are naturally less commonly addressed.

The Illustrations of the IPA series set the stage for greater acoustic detail in phonetic descriptions by exemplifying the use of the much-expanded post-Kiel IPA. Even though coverage through the first decade of sketches was sparse in phonetic detail by modern standards, the Illustrations series marks a significant step forward in phonetic documentation, using the IPA as a common character set that allowed better comparisons across languages. Integrating phonetic measures that have since become more commonplace was a natural progression. Transcription has always been acknowledged to have limitations, of course, but it is often the case that those limitations are ignored once the transcription is available (e.g. Errington 2007: 8–9). Greater mismatches are felt by many current approaches to phonetics and phonology or in cases of ‘covert contrasts’ (Munson et al. 2010). The range of variation that exists cannot be captured if the transcriptions do not represent it, and current approaches are more likely to take the variation as necessary for a full understanding of the phonetics of a language (e.g. Hay & Drager 2007; Ladd 2014).

The boundary between documentation and theoretical research is ill-defined, and it always will be. We can expect that the boundary will shift toward including more aspects as ‘documentation’ as the relevant measures become more agreed upon, but there will remain aspects of phonetics that will only yield to experimentation and thus should not be expected to be part of a documentation effort. Most importantly, the categories need to be hierarchically arranged, so that aspects of each category (such as duration) can be seen separately (rather than being separated for some categories (e.g. vowels) and not others (e.g. resonants)). The numerical scoring needs further improvement as well. We have already mentioned that it is
unfair to penalize a documentation article for not describing a non-existent aspect of a language (e.g. implosives or pre-aspirates). Even stronger variants of the numerical assessments could be obtained by weighting them by the number of talkers and the number of tokens per talker. We believe that concerted effort in standardizing the categories, and thus promoting consistency in phonetic documentation, is the only way to move forward.

Perception of phonetic aspects was excluded from our survey, but that should not blind us to its importance and urgency. We can measure every millisecond of every recording ever made of a language and still not be sure whether the native speakers paid attention to what we measured. For example, formant frequencies generally covary with distinctive vowel length, but speakers of Japanese, for example, do not use that information perceptually (Lehnert-LeHouillier 2007). When languages become endangered, the remaining speakers can generally produce the phonemes and texts they learned from childhood onwards, but decrements may occur due to misremembering, hearing loss, and interference from a later-learned but more commonly used language. Therefore, assessing the perceptual value of phonetic properties in endangered languages with some younger speakers is, in some ways, more urgent than the documentation of other aspects. Perception, however, also tends to be on the ‘theoretical research’ side of the theoretical research/documentation continuum, and the number of perceptual effects that have been found is quite large, making choices about which perceptual tests to run highly problematic. Several of the surveyed papers in *JPhon* (Beddor, Harnsberger & Lindemann 2002, Gerfen & Baker 2005, Kirby 2014) did indeed perform perception studies to address issues raised by their measurements. Other *JPhon* papers that were excluded (Harnsberger et al. 2001, Lehnert-LeHouillier 2007, DiCanio 2012) were primarily concerned with perception. (The DiCanio article’s acoustic measurements were for coarticulation, which we excluded from our list of documented effects, as noted above). Because perceptual studies have not resulted in a standard set of measures (as acoustic measures barely have), it may be some time before consensus on the tests that would be needed for perceptual documentation can be agreed upon. However, the need exists, and, for many languages, it is urgent.

A similar set of concerns applies to voice quality and to physiological measures. Some acoustic measurements of voice quality have had fairly wide-spread use (see the studies reviewed in Kreiman & Sidtis 2011), but they have generally been employed to study linguistically distinctive use of voice quality rather than making an assessment of voice quality across all languages. Should those measurements be made more generally? Doing so could add valuable insight into previously unexamined aspects of languages, but it could also add a daunting amount of work to what is already a labor-intensive process of documentation. Physiological measures have also generally been used for relatively narrow theoretical questions, but measurement devices are becoming more widely available and portable. At what point do these improvements in efficiency merit their inclusion in a phonetic documentation standard?

Two aspects of speech that have received a great deal of attention in *JPhon*, coarticulation and domain-initial strengthening (e.g. Cho & Keating 2001), were excluded from this survey as being still within the realm of theoretical research rather than documentation. Given their importance and universality, it is desirable that they be included in documentation at some point. This would require decisions about how to quantify both domains; this has yet to be accomplished. Again, this is an aspect of documentation that will need elaboration in the future, as theoretical research methods move into the documentation domain.

The recent increase in the rate at which languages are falling silent due to loss of speakers has led many communities to establish revitalization efforts of many sorts (e.g. Hinton 1994, Hinton & Hale 2001, Hinton, Huss & Roche 2018). Some of these efforts entail recovering languages that have lost the continuity of typical language transmission, thus requiring the use of historical records and comparative methods (Amery 1995, Baldwin & Olds 2007, little doe baird 2013). As Bird & Kell (2017: 539) point out, ‘[m]ost Indigenous language
revitalization programs in Canada . . . currently emphasize spoken language. However, we still know very little about second language (L2) learning in the context of Indigenous language revitalization . . . particularly with respect to pronunciation’. Nonetheless, phonetics is seldom mentioned in the revitalization literature. One collection contains almost no mention of phonetics (Coronel-Molina & McCarty 2016). Another mentions phonetics as one of the gaps in the written representation which would be augmented somewhat by audio recordings (Spence 2018: 183). Even audio recordings, while immensely valuable to language learners, generate sounds that are not produced by speakers in the flesh, leading one language activist to label them products of ‘zombie linguistics’ (Perley 2012). A few projects have examined the usefulness of using phonetic material other than just audio in the service of language revitalization. A study of Cherokee tone, with emphasis on the usefulness of the analysis to language maintenance, has appeared (Herrick et al. 2015). Visualization via ultrasound has proven effective in improving pronunciation in indigenous languages (Bliss et al. 2018). The success of these studies indicates that better phonetic documentation would be useful in devising aids for revitalization programs.

One surprising aspect of this survey was the extent to which English-language research focusing on the phonetics of Indo-European languages (including multiple dialects of languages like English) has continued to comprise an equivalent percentage of descriptive phonetic work over the decades. Despite several funding initiatives (e.g. the Hans Rausing Endangered Language Project, the U.S. National Science Foundation’s Documenting Endangered Languages (now Dynamic Language Infrastructure) program, and the Endangered Language Fund) that have sought to highlight research on endangered and minority languages, the phonetics of most of the world’s languages remain vastly unexplored. This bias in language area distorts not only which aspects of speech and their patterns that we consider worthy of investigating in greater detail, but it also underreports the true range of diversity in human language. For instance, many languages lacking phonetic descriptions altogether have substantial dialectal diversity, but given the number of studies focusing on English phonetics, a reader might falsely conclude that variation in English is inherently more relevant for phonetic analysis. This survey suggests that greater phonetic research on a variety of language families is sorely needed.

5 Next steps

No language has ever been fully documented, and it seems impossible for that to happen. Guidelines for doing documentation do exist, of course, acknowledging that good documentation is better than perfect-but-nonexistent documentation (Ladefoged 2003, Woodbury 2003, Bhaskararao 2004). The categories chosen here were based largely on aspects covered in the Illustrations of the IPA. We do not claim that these are the best categories; they are focused on acoustics, while articulation and perception are important and increasingly-addressed aspects as well. We believe that a larger survey of phoneticians as to what would constitute nearly complete documentation is in order. Future surveys should include research published in languages other than English, of which there are many. Of particular interest will be phonetic measures that are prominent in non-English publications that are rare or absent in English-language ones. In any literature, features tend to begin as theoretical research topics, and, once they are better understood, they can move into documentation; this process is bound to continue, so the list will never be static. In addition, it would be desirable to have a better algorithm for calculating amount of coverage than the one we used here. A report should not be penalized for not measuring an aspect that the language lacks, such as preaspiration or tone; our rough, first pass does so. The algorithm should presumably be weighted so that the number of speakers reported is taken into account. We did attempt to determine the number of speakers consulted from the Illustrations series, but often it is not explicitly
reported, or reported in vague terms. Nearly half of all Illustrations (71 of 161, 44%) report data from a single speaker. Having an inventory and addressing the issue of the number of speakers would, we believe, be of use to researchers who might not have thought to measure certain aspects of a language that they could, with greater or lesser effort, measure. For example, most surveys neglect sonorants altogether, even though they could be measured. Even if completeness is unattainable, knowing how close a description comes to completeness is worthwhile. We hope that, in the not too distant future, a catalog of desirable measurements will be agreed upon.

Acknowledgements

This work was supported by NSF-DEL grant #BCS 0966411 to Haskins Laboratories and NIH grant R01 DC-002717 to Haskins Laboratories. Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the National Science Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank the editors, three anonymous reviewers, Laura Koenig, Christine Shadle and Rory Turnbull for helpful comments.

Appendix. Articles in each collection

ILLUSTRATIONS OF THE IPA (JIPA)


Phonetic documentation in three collections: Topics and evolution


SOUNDS OF THE WORLD’S LANGUAGES (SOWL) DOCUMENTATION PROJECT


---

**References**


