ON THE STRUCTURE OF VOWEL SPACE: A GENEALOGY OF GENERAL PHONETIC CONCEPTS

Hendrik F. V. Boshoff and Elizabeth C. Botha

Department of Electrical and Electronic Engineering, University of Pretoria, Pretoria, 0002, South Africa

Email: hendrik@suntiger.ee.up.ac.za

ABSTRACT

This paper gives an overview and an analysis of representative vowel space depictions found in the phonetic literature of the past four centuries. A visual representation of 24 selected structural proposals with time is given in poster form, after translation to a common notation.

A first-order approach is developed by which the distance between any pair of vowel schemes may be measured. Using these distances and other considerations, trends are detected and major innovations identified. An attempt is then made to trace the genealogy of five general phonetic concepts concerning the structure of vowel space: discretizing the continuum, the assumption of symmetry, triangular vs. square arrangements, the development of 'roundedness' as an independent dimension and the treatment of 'central' vowels.

Such a historical perspective is deemed necessary for a full appreciation and utilization of the current vowel schemes.

1. INTRODUCTION

Machine processing of spoken language always involves the transition from continuous acoustic signals to discrete linguistic symbols. This operation constitutes a form of phonetic and/or orthographic transcription. The finite set of symbols chosen as the basis for such transcription reflects fundamental theoretical considerations, as well as many historical influences. Without necessarily being an expert phonetician, the worker in this interdisciplinary field is often faced with important phonetic decisions, especially when selecting a notation.

One example of this is the choice between a broad and a narrow transcription. Broad transcriptions omit some details, which may be systematically implied, such as the glottal stop before word-initial vowels in some languages [1]. These omissions have to be compensated for when using the transcription. Although it is widely used, broad transcription practices are aimed at speeding up and facilitating use by people, and may not be the most suitable for machine implementation of any (even speaker-independent) task.

The production mechanism of vowels allows a more continuous transition between them than between most consonants. The general phonetic problem of the structure and discretization of vowel space has therefore generated much controversy, and current vowel schemes have to be viewed in historical perspective to judge their merits, and to use them properly.

Vowel symbols appear in a set sequence in every alphabet. Different linear orderings based on physiological and auditory considerations were proposed by Robinson (1617), Olearius (1630), Newton (1665) and Kempelen (1791) [2,3].

The first tabular arrangement of vowels was made by Brightland in 1711; his work may be regarded largely as a translation of Wallis's 1653 book *Grammatica Linguae Anglicanae* [2]. Hellwag's triangle of 1780-81 is also two-dimensional [4], but rather than being a table, it may be viewed as an explicit attempt to suitably discretize the continuum of vowel space by placing vowels in a somewhat freer geometric arrangement.

The second half of the 19th century was a particularly fertile period for extending and augmenting earlier vowel schemes. Brücke, Winteler, Sievers and Trautmann were among the users of the triangular scheme [5]. Melville Bell reinvented the 3x3 square table in 1867, and increased the number of vowel space dimensions to four. His approach was extended by Sweet [6].

Passy's attempt at a compromise between the competing ('triangular' and 'square') schools [7] largely influenced Jones [8] and the schemes eventually adopted by the IPA [1]. With the ascendancy of the IPA, the need for standardization was recognized and widely respected. However, examination of the record shows that consensus had not yet been reached on the entire structure. Particularly the area around the vowel [a] was still subject to different representations. It is probably no coincidence that more than one would-be reformer of the IPA vowel scheme targeted the neighbourhood of [a] in this century.

Recognizing that the widely used IPA (1993) vowel scheme is a somewhat unsatisfactory combination of two historically separate approaches prompts one to question its optimality. Given its entrenchment however, any proposed structural modification would have to be very well motivated. A historical perspective is essential for such an exercise, but also for a full appreciation of the status of any selected vowel notation.

In the next paragraph we describe our selection and display of data, while a number of general phonetic concepts are introduced in §3. Our comparison of the selected vowels schemes is presented in §4, followed by the tracing of the genealogy of several concepts, and a conclusion.

2. DATA SELECTION AND DISPLAY

A selection had to be made of vowel schemes to be included in this study. Only the later ones were available to us in the form of original publications by their authors. Reference was therefore made to secondary sources [2,3,5,6,9,10], which were cross-checked as far as possible. We also deliberately limited ourselves to the use of diagrams only, most of which had been taken directly and completely from the original sources. These 'skeletons from the fossil record' were used without regard to any motivating and sometimes controversial considerations employed during their construction, whether auditory, articulatory or acoustical. Diagrams that were structurally identical or very close to earlier ones were excluded, as were some of the more complicated ones (e.g. Evans' of 1882 [5]).

The 24 selected diagrams are available in their original form on the site [http://www.ee.up.ac.za/hendrik/icslp5/diagrams.htm]. As could be expected, the notation varies widely. The first step towards comparison is conversion to a uniform notation. The poster shows all the diagrams, converted to IPA notation, arranged along a time line (CD-ROM [0969.jpg]).

3. GENERAL PHONETIC CONCEPTS CONCERNING VOWELS

In this paragraph a number of general phonetic concepts concerning vowels will be introduced and briefly discussed, as they form the basis of our comparison of vowels schemes.

Discreteness and continuity. The problem of discretizing the continuous speech signal has already been mentioned. In addition to this, vowel space itself is a continuum that is usually discretized. All the schemes considered in this paper use a finite number of symbols to label regions or points in vowel space. The number of discrete levels into which each dimension is divided may vary from two to seven. Stevens in particular drew attention to what he called the 'quantal nature' of speech [11].

Symmetry assumptions. The earliest 2-dimensional diagrams and tables all assumed a perfect symmetry, at least in the 'front-back' dimension. A case may even be made that vowel systems in early and present alphabets exhibit this symmetry [cf. 12]. No explicit argument could be found to support this assumption, although it is almost universally made.

The 'affinity structure' of vowel space. Neary introduced the idea of an affinity structure of vowel feature systems as an abstract set of 'relationships of partial identity and relative proximity that are assumed to hold among a set of phonetic segments' [13]. We may postulate that some notion of this relative similarity between vowels is what phoneticians try to clarify when they draw vowel diagrams.

The dimensions of vowel space. Vowels have been classified along any number of dimensions between one and seven. The articulatory system prevailing at present [14] uses three, which may be abbreviated to 'high/low,' 'front/back' and 'rounded /unrounded.' Although these terms are used, they will be treated as empty labels, and not referring to tongue and lip positions as originally intended.

The shape of the vowel figure. Not everyone making some representation of vowel space drew an outline of the area where the vowels had been placed. However, the region is universally considered to be bounded, and could be described as roughly either square, triangular or quadrilateral in many cases. A stylized figure is sometimes used for ease in drawing (e.g. [14]).

Regularity of reference point placement. A quick inspection shows that the labeled qualities are very evenly distributed in all the schemes selected for this study. In the case of the cardinal vowels this question is discussed by Jones [8], who points out that his front cardinals are slightly further apart than the back ones.

4. COMPARING VOWEL SCHEMES

There is no obvious way to compare geometric arrangements ranging in dimension form one to five and using between seven and 72 discrete elements. (We did not use a graph theoretical approach, which might be fruitful.) All general vowel schemes do include the three 'point' vowels [i,a,u] however, and arguably relate to the same 'substrate' of speech sounds. It is deemed desirable to be able to classify the structures in some way, and even to define a distance measure between any two of them.

4.1 Structural Characterization

The number of dimensions used in arranging vowels is clearly important for the structural characterization of the scheme. The number of labeled qualities is of secondary importance, as they are often taken to be divisions of arbitrary size. The symmetry of the pattern should also come into play, and once an identity can be established between symbols used in separate schemes, the adjacency implied by their relative placement as well.

We found that a simple *structural formula* could be used to capture at least some of these factors. The discrete number of labels of 'height' (or 'openness') is first determined, and the number of labeled qualities at each level counted. These counts are then arranged starting at the level of [a] and ending at the level shared by [i] and [u]. For example, the original triangular scheme of Hellwag has the structural formula 1-2-3-3. When an explicit separation along binary features is part of the diagram, such sub-diagrams are assigned formulas separately. Bell's table of 1867 has four identical parts, and can hence be described by the formula 4x(3-3-3). The poster shows the structural formula for every vowel scheme.

4.2 Relationships between schemes

Being a discrete placement of symbols in a continuous space, one vowel scheme may be transformed into another using the following operations:

- Insertion and deletion of labeled qualities
- Geometric rearrangement with conservation of topology
- Changing the topology with dimensions conserved
- Altering the number of dimensions

Symbols added to connecting lines on the poster indicate which operations would transform the starting to the ending diagram. The fewer transformations required, the closer can two schemes be considered to be. Definition of a distance measure requires quantifying this closeness, which in turn implies the combination of relevant operations. The main difficulty with such combination lies in the relative weight assigned to different operations, e.g. the addition of a quality vs. the addition of a dimension.

A distance measure: In order to get started, the following first-order approach is proposed.

- A *reference vowel structure* is defined, having the three primary dimensions of vowel quality mentioned above. Five levels of discretization is assumed in the 'height' dimension and two in 'roundedness,' with 'frontness' left undetermined.
- Each diagram (except the linear ones) is mapped to the reference structure by separation of 'rounded' from 'unrounded' vowels and insertion or deletion of levels in 'height' and other dimensions if necessary. Track is kept of the mapping details.
- All such schemes can then be assigned two five-number structural formulae, as defined before.
- The *reference structure* distance between two schemes is determined by adding together the absolute value of differences in corresponding numbers of both structural formulae. The *mapping* distance is taken simply as the number of operations necessary in mapping both schemes to the reference structure, with operations in common eliminated first.
- The *total* distance between two schemes is a weighted combination of the reference structure and mapping distances in an empirically determined ratio of 1:10. The total distance is also normalized by the geometric mean of the number of labeled vowel qualities in each scheme.

This definition is meant to yield a measure meeting the formal requirements of a metric, but no attempt will be made to prove that. It should be intuitively clear at least that the measure is positive definite and symmetric, and that the self-distance is zero for all cases. Linear schemes are not included, because the mapping step would require too many arbitrary allocations.

Trends: Using the proposed measure, a distance matrix may be calculated for all the selected vowel schemes (see the poster). It immediately shows that some distances (e.g. between Chladni 1809 and Trautmann 1877 [5]) are zero, implying that they share the same affinity structure under this measure. Bell's scheme on the other hand is remote from nearly all the others, partly due to the large number of labeled qualities. Jones's cardinal vowel system was measured to be very close to Passy's second scheme, as could be expected from examining the diagrams.

The numbers in the distance matrix can be further analyzed in many ways. One approach builds on the observation that an original and influential scheme will be closer to later diagrams than to those of its own or earlier times.

If we make the simplifying assumption that each selected scheme advances the innovation clock by one step, the columns of the distance matrix can be viewed as a series of distance measurements uniformly spaced in (warped) time. It then becomes possible to determine a tendency by fitting a straight line to the numbers in each column. Lines with a negative slope indicate 'forward looking' proposals, while those with a positive slope agree better with earlier times.

A plot of the calculated slopes is shown in Figure 1. The slopes are normalized with the mean distance value; otherwise a scheme that is far removed from all the others could have a misleadingly dominating slope.

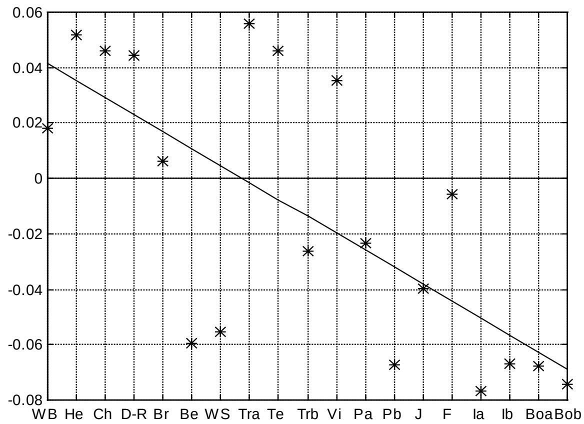

Figure 1: Normalized slopes of columns from the distance matrix, arranged chronologically. Positive values indicate schemes closer to those of early times, and negative ones were closer to later times. A trend line is added to show clearly which schemes lead their own time (below the line) or lagged (above).

The linear trend shown is not meant to reflect a positivistic attitude. We do not want to assume that the trend approaches a 'true' value and the selection of schemes could also influence the exact values obtained. What can be said is that the consensus apparently shifted over the past few centuries, and that major innovators (and laggards) clearly stand out in this representation.

The question could be asked what constitutes the innovations of Bell, Winteler, Sievers, and Passy, which set them apart. In the next paragraph an attempt will be made to trace the genealogy of some concepts.

5. TRACING THE GENEALOGY OF SPECIFIC CONCEPTS

Some of the innovations in vowel diagramming developed gradually, and were contributed to by a number of authors. Others were introduced in a radical move by a single worker. Five concepts were selected for closer scrutiny.

Discretization and rediscretization: Early on a small number of discrete vowel qualities were depicted, only slightly extending the meager symbol set of the Roman alphabet, and at first adding the extra qualities found symbolized in the Greek alphabet. A gradual growth in number appears in the 'triangle' tradition, from Hellwag's nine to 18 with Sievers [5]. Melville Bell's 36 vowels [6] were at least twice as many as those of any of his contemporaries, but he was partly constrained by his simplistic doubling of qualities when adding a dimension. The present view seems to be that 36 are too many, and few phoneticians take Sweet's extension to 72 seriously today. The IPA (1993) uses 28.

Contrary to the later tendency, Du Bois-Reymond significantly *decreased* the number from Hellwag's nine to seven in 1812 [5]. Perhaps this may be attributed to an analogy with the colour-mixing triangle of Mayer (cf. [9]), a view which was advocated even in this century [3]! The seven positions correspond with the simplest interpolation between three corners of a triangle.

Symmetry and its breaking: The first deviation from 'front /back' symmetry seems to appear only with Viëtor in 1884. His mild skewing of the diagram is based on articulatory considerations, and a similar deformation appears later in the cardinal vowel diagrams of Jones and the IPA. In all these cases the number of discrete levels are still the same in the front and at the back however.

Square vs. triangular arrangements: It may be observed firstly that the appearance of planar diagrams promptly lead to the extinction of linear schemes. The first square table predated the first triangular diagram by nearly 70 years, but it is not clear whether Bell was influenced directly by Brightland. For most of the 19th century, the two shapes were in opposition. It was the Frenchman Passy who first attempted the uneasy marriage of the 'German' triangle to the 'English' square in 1890 [7].

Most 20th century schemes clearly carry features of both. Present day opposition to insertion of an 'open central' vowel (e.g. [a]) into the IPA diagram, can be construed as 'triangular' resistance to the complete 'squaring' of this scheme. And most of the awkwardly placed symbols in the 1993 diagram are carry-overs from the 'wide' vowels of Bell and Sweet.

Development of 'roundedness' as independent dimension: In the days before Bell, 'rounded' and 'unrounded' vowels were found side by side in the same plane. Brücke [5] introduced two 'unrounded back' vowels to symmetrically balance the 'rounded front' vowels in his scheme of 1856, but a full complement of both didn't fit well with the triangular geometry. Passy kept unpaired 'open' vowels as late as 1890. Bell completed the pairing in 1867 by placing 'rounded' and 'unrounded' vowels in two separate planes.

This independence of the 'roundedness' dimension influenced the later triangles, which morphed into 'stars' with Sievers and Trautmann, without changing their structural formulae. A late example of incomplete acceptance of full pairing is furnished by Jones's initial reluctance to include [OE] in his scheme.

Even though the IPA (1993) again represents 'rounded' and 'unrounded' vowels together on the same plane, the conceptual separation is complete.

Treatment of 'central' vowels: While Bell only *concluded* the 'roundedness' pairing, he alone must be credited with the introduction of the 'central' vowels (which he called 'mixed'). In the triangular diagrams, the central positions were already occupied. Passy later forced the way open for 'central' vowels, but sacrificed one of Bell's six to the triangle principle. This one was eventually restored, along with two 'rounded open' vowels by Jones and the IPA.

The 'mid central' or 'neutral' vowel *schwa* was taken by Jones to contrast with all his cardinal qualities, which he considered to be *peripheral*. In this sense it forms the hub of his system.

6. CONCLUSION

An important result of this study, is the extent to which the widely accepted IPA (1993) vowel scheme is an unsatisfactory compromise between two historically separate approaches. No amount of tinkering within the present framework will rectify this. Structural modifications will have to address the general phonetic issues raised in this study. One proposal based on a scaling discretization of formant space (SDF) [15] is included in the poster. Its distance to the IPA scheme is relatively small, and it rectifies some of the perceived shortcomings.

7. REFERENCES

1. Principles of the IPA, London, 1949 (reprinted 1964).
2. Ladefoged, P., *Three Areas of Experimental Phonetics*, OUP, London, 1967.
3. Russel, G.O., *The Vowel*, OSUP, Columbus, 1928.
4. Jakobson, R. and Waugh L. *The Sound Shape of Language,* Mouton, 's-Gravenhage, 1987.
5. Viëtor, W., *Elemente der Phonetik*, 7. Auflage, O.R.Reisland, Leipzig, 1923, (1. Aufl. 1884).
6. Henderson, E.J.A., *The Indispensable Foundation*, OUP, London, 1971.
7. Passy, P., *Changements Phonétiques*, Paris, 1891.
8. Jones, D., *An Outline of English Phonetics*, 9th Ed W.Heffer & Sons, Cambridge, 1964, (1st Ed. 1918).
9. Zwirner E. and Zwirner K., *Grundfragen der Phonometrischen Linguistik*, 3. Aufl. Karger, Basel, 1982.
10. Panconcelli-Calzia, G., *Geschichtszahlen der Phonetik* (1941) and *Quellenatlas der Phonetik*, (1940) John Benjamins, Amsterdam, 1994, (K. Koerner, Editor).
11. Stevens, K.N., "On the Quantal Nature of Speech", J. of Phonetics 17: 3-45, 1989.
12. Miller, D.G., *Ancient Scripts and Phonological Knowledge*, John Benjamins, Amsterdam, 1994.
13. Neary, T.M., *Phonetic Feature Systems for Vowels*, Indiana Linguistics Club, Bloomington, 1978.
14. Abercrombie, D., *Elements of General Phonetics*, Edinburgh UP, Edinburgh, 1967.
15. Boshoff, H.F.V. and Botha E.C., "An acoustical analysis of the cardinal vowels as spoken by Daniel Jones," IEEE South African COMSIG Conference, 1998, to appear.