iPA Phonetics: MULTIMODAL iOS APPLICATION FOR PHONETICS INSTRUCTION AND PRACTICE

John H. Esling[1], Scott R. Moisik[1+2], Christopher Coey[1+3]

[1]University of Victoria, [2]Max Planck Institute for Psycholinguistics, Nijmegen, [3]CD-Mediaworks, Victoria, BC
esling@uvic.ca, Scott.Moisik@mpi.nl, coey@uvic.ca

ABSTRACT

iPA Phonetics is an iOS application based on the Laryngeal Articulator Model that illustrates voice qualities, vowel qualities, and consonant production with video/audio of articulations in the oral vocal tract and laryngoscopic video/audio and ultrasound images of the laryngeal vocal tract. The Consonant Chart is an expanded version of the 2005 IPA chart; the replay speed of consonant and vowel sounds can be controlled; and random consonants and vowels can be generated for recognition practice. This self-contained app for Apple iOS mobile electronic devices gives users the ability to access and compare phonetic symbols and sounds together with detailed articulatory production correlates. The app is free on the Apple Store, intended to introduce the general public as well as specialized users of IPA symbolization, via iPad/iPhone technology, to the auditory inventory of possible speech sounds of the languages of the world and to how each sound is physically articulated.

Keywords: application, multimedia, laryngeal, instruction, articulation.

1. INTRODUCTION

iPA Phonetics is a new iOS application (app) that illustrates the sounds and articulations of a range of phonetic categories, including an expanded version of the IPA chart [1]. This app gives users of Apple iOS mobile electronic devices the ability to access, see/hear, compare, and test their knowledge of phonetic symbols and sounds together with their visual production characteristics, including video of the oral vocal tract, laryngoscopic video of the laryngeal vocal tract, and ultrasound of the laryngeal vocal tract. There are four pages that can be selected in the app: Consonants, Vowels, Voice Qualities, and an About (information) page that gives references to the background and origins of the app and to IPA resources that are relevant to using the app. The Voice Qualities page contains a diagram of the vocal tract that follows the Laryngeal Articulator Model (LAM) [5, 3, 4], which can be used as a reference for the "auditory location" of the parallel places (columns) and manners (rows) of articulation of consonants in the Consonant Chart or of the relationship to vowel qualities in the Vowel quadrilateral. The app is free and entirely self-contained, in that once downloaded it requires no internet connection to use.

The purpose of the app is to serve as a resource to introduce linguists, students of IPA symbolization, and the general public via iPad/iPhone technology to the auditory inventory of possible speech sounds of the languages of the world and to how each sound is physically articulated. It is a learning tool created in a university environment for the benefit of students, teachers, researchers, and wider communities that work with language. The taxonomy is also a window on the theory of phonetic classification.

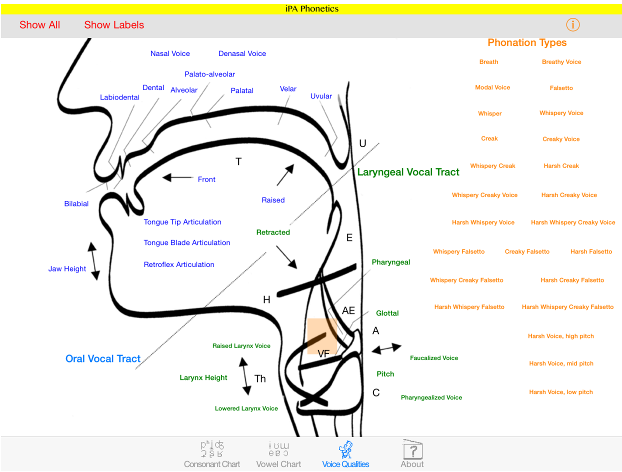
2. STRUCTURE OF THE APP

2.1. Components and functionality

The Consonant Chart format of the opening page follows the elaborated table for phonetic notation of speech sounds in *The Handbook of Phonetic Sciences* [6]. The Vowel Chart of the second page follows 2005 IPA presentation. Each symbol in either chart is linked by touch to an accompanying video/audio file (or set of files). Views of consonant and vowel categories are presented in pop-up windows that appear in the lower-left corner of the screen. Each consonant or vowel category can be listened to and viewed in the form of close-up oral-endoscopic videos of the vocal tract. Images/audio may be sped up or slowed down to compare the degree of auditory distance between categories. A "Playback Controls" option can be selected to stop, advance or review the video, or to expand the pop-up window to full screen.

The app has a Voice Qualities page, shown in Fig. 1, with a static "map" of the vocal tract overlaid with geographically oriented labels for long-term oral and laryngeal voice quality settings, each touch of which plays an audio file of the corresponding oral or laryngeal category. Key regions are also depicted, which contain subsets of the inventory, organized into groups according to the LAM theory of where each quality is generated articulatorily. These include Jaw Height, Larynx Height, (Tongue) Front, (Tongue) Raised, (Tongue) Retracted, Glottal, and Pitch regions, as well as a region within the epilaryngeal tube that is a touch pad for an array of phonation types, which are also listed individually at the right of the page. Each touch therefore plays one designated category or one in a rotation of related voice quality tokens, some of which overlap across regions, depending on the complexity of the region, the epilaryngeal tube "box" being the most complex.

Figure 1: Voice Qualities page of *iPA Phonetics*.

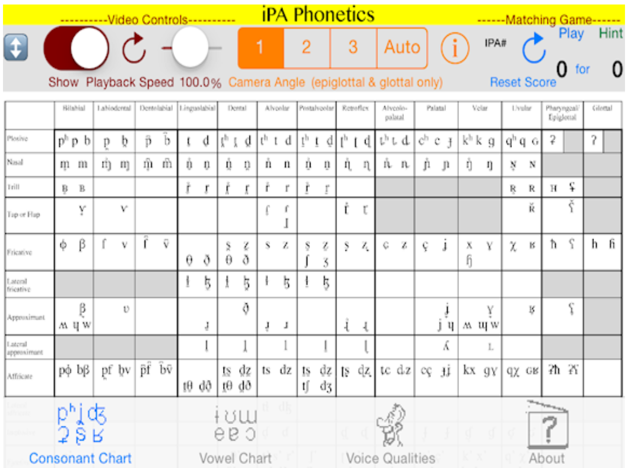


2.2. Design background and features

The Consonant Chart (Fig. 2) contains 14 columns and 13 rows, adding some categories to the standard set of IPA places and manners, drawing on the ExtIPA chart [8]. These are categories that might be considered "normal phonetic" possibilities rather than primarily disordered/clinical, i.e. the columns Dentolabial and Linguolabial. An Alveolo-palatal column is added, expanded from its brief mention in the official IPA chart. Pharyngeal is identical in articulatory stricture location to Epiglottal [5, 6, 11]. There are four major divisions in the Consonant Chart (separated by bold lines), which are intended to guide the user to identify groupings of major articulator movements. The first three columns are Labial; the second five columns, Front Lingual; the third four columns, Back Lingual; and the last two columns, Laryngeal. Symbols are included for as many cells in the matrix as phonetically practicable, given articulatory possibilities. Shaded areas denote phonetic impossibilities. Diacritic placement follows IPA principles. The Plosive row explicitly includes aspirated stops, especially where their linguistic incidence is more common. The Nasal row is fully elaborated for all places with a voiceless counterpart in each cell. Trills involve the lips, tongue tip, uvula, and the aryepiglottic folds of the Pharyngeal/ Epiglottal articulator (using the IPA symbols for "Epiglottal fricatives" [6]). The Labiodental/Bilabial flaps are distinguished; a Uvular tap (controversial in past IPA charts) is given a diacritic; and a newly designated Pharyngeal tap appears [4, 6]. Fricatives are divided into sub-rows where multiple tongue configurations can occur at the same place, e.g. due to grooving or sibilance. The Approximant row includes a sub-row of "approximated fricatives" [9]. Affricates are elaborated, in parallel to Fricatives, including two Pharyngeals. Implosives, Ejectives (with affricates, in parallel to the Affricate row), and Clicks (the sparse set of 5 from the IPA chart rather than a full set as in [10]) are added as a manner row.

Laryngeal articulations are accompanied by a choice of three laryngeal ultrasound images as well as laryngoscopic videos of the deep vocal tract.

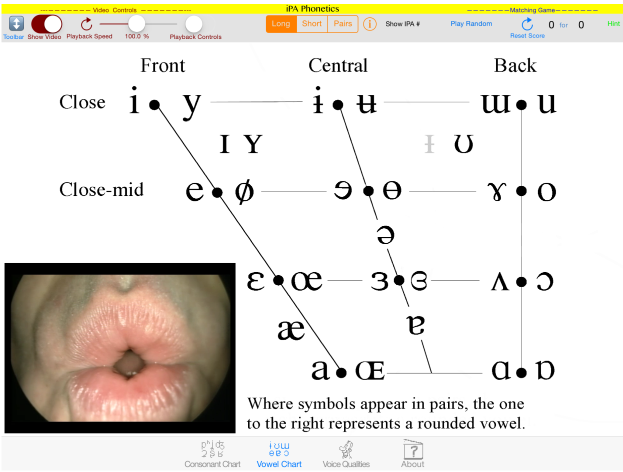
Figure 2: The main page of the *iPA Phonetics* app, showing the top portion of the Consonant Chart, option selection controls, and links to the Vowel Chart, Voice Qualities and About (credits) pages.



The Vowel page consists of a Vowel Chart after the 2005 IPA chart, with one added symbol for a near-close near-back unrounded vowel, completing a vowel pair that is worth elaborating on the chart. This page, with selection options in the top bar, is illustrated in Fig. 3. Vowel sounds can be played as long, short, paired (two at a time), or "swept," i.e. a sequential auditory/articulatory continuum from one corner or side of the vowel chart to another – a novel programming feature for phonetics illustrations that takes advantage of iOS capabilities. Recognition can be tested via a "Play" button that randomly presents consonant or vowel clips and keeps score.

Neither the Vowel nor Consonant Chart additions are meant to replace the 2005 IPA chart as the official inventory of officially recognized symbolic representations for sounds. The symbols available in the app, and their organization, are meant to be a logical extension and economical reorganization of the symbol categories in the 2005 IPA chart. This reorganization is due in part to the capability introduced by iOS screen displays. Because the table can be expanded and moved around on the screen using 2-finger control, increasing the number of columns and filling in more cells and rows does not pose the visual clutter problem that it would in a static printed display. The official IPA Number sequence of each symbol can also be toggled on/off (Show IPA#) and viewed beneath each symbol.

Figure 3: The Vowel Chart page of the *iPA Phonetics* app, showing a pop-up window open, playing a rounded vowel.



3. TECHNICAL CONSTRUCTION

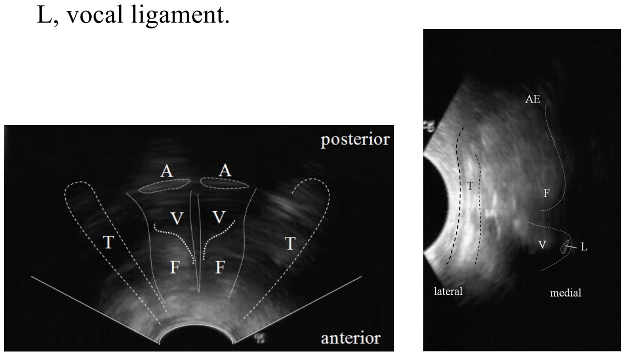
Video/audio clips for each set of symbols were captured in a sound-booth environment with a high-quality microphone and using a KayPENTAX 9100 Rhino-Laryngeal-Stroboscope, 9105 70°-angle rigid oral laryngoscope (hand-held), a 35mm lens to view oral articulations and a 28mm wide-angle lens for laryngeal articulations, connected via a Panasonic GP-US522 camera to custom-designed multichannel synchronizing software. Oral consonants and vowels were filmed with the rigid endoscope rod held in front of the mouth. Pharyngeal/Epiglottal and Glottal consonants were filmed with the rigid endoscope rod inserted into the mouth to view the laryngeal area as is the practice in laryngoscopic experiments [3]. This allows the vocal, ventricular, and aryepiglottic folds, the epiglottis, epilaryngeal tube in general, and larynx height characteristics to be viewed during the production of these otherwise difficult-to-observe articulations. This is critical in phonetics instruction, as these structures have been shown to be responsible for the generation of multiple types of periodic vibration and complex modification of the pharyngeal resonating chamber, accounting for a wide range of contrastive auditory qualities in the languages of the world [4].

Three different ultrasound images were also captured for each of the 12 sounds in the Laryngeal (Pharyngeal/Epiglottal and Glottal) consonant range to give a clearer picture of the changes in laryngeal constriction and larynx height that occur during the production of these deepest articulations. Ultrasound images of the laryngeal part of the vocal tract have not previously been included in databases of articulatory categories (cf. the lingual ultrasound images for oral sounds in [12]).

Ultrasound images that accompany Laryngeal articulations were captured using a GE portable LOGIQe R5.0.1 system with an 8C-RS probe to image supraglottic laryngeal involvement (e.g. for Laryngeal stops) and with an 12L-RS straight-line probe at a relatively shallow 2-4 cm depth on the neck and 2-4 cm of the vertical dimension for clear resolution of laryngeal structures to image larynx height changes (e.g. for Pharyngeal/Epiglottals). This is a novel laryngeal technique (see [11]) that differs from the approach usually taken in oral lingual ultrasound data capture.

The three ultrasound viewing options are (1) an axial convex view (Fig. 4, left), (2) a vertical convex view (Fig. 4, right), and (3) a vertical flat (straight-line) view showing up-and-down neck muscle movement. Anatomical sections of larynx models are included with corresponding labelled ultrasound images in a supplementary "i" = information page, viewable via the toolbar, to illustrate the detail of what can be seen in the laryngeal ultrasound.

Figure 4: Axial convex (left) and Vertical convex ultrasound (right). A, arytenoids; V, vocal folds; F, ventricular folds; T, thyroid cartilage; AE, aryepiglottic folds; L, vocal ligament.



4. THEORETICAL FOUNDATIONS

4.1. Why the Laryngeal Articulator Model?

It was considered that the LAM would give users, especially learners of phonetics, a more grounded perspective on the concept of voice quality (deriving not only from a voice source but also multi-layer articulatory mechanisms), on the consequent relation between voice quality postures (and their auditory correlates) and vowel quality, and on how the Laryngeal columns of the Consonant Chart relate to voice quality settings [3]. This is in contrast to a lingual model that presents only a consonant or vowel inventory, assuming the larynx contributes only voicing to the sound stream while all other filtering is a function of tongue shaping from larynx to lips. The LAM has implications for theories of speech acquisition and sound change, since the "Laryngeal Articulator" is the principal mechanism that infants first learn to control as they test and practice their phonetic production skills from birth through the first several months of life [2, 7].

4.2. The app as instructional resource: understanding phonetic categories through voice quality

From a learning point of view, it was therefore felt that the auditory/acoustic cues that can be heard in the app for consonants, vowels, or voice quality can be associated by the listener with particular sounds occurring in the wide range of languages they may be familiar with. Drawing on their prior linguistic experience, listeners may also associate these three levels of sound types in the app with elements of sound production that they have observed in early infant vocalizations. The design of the app is thus heavily weighted towards the descriptive elaboration of consonants in the Laryngeal sector and their inherent connection with categories that are grouped on the Voice Qualities page in the lower, laryngeal half of the vocal tract. The intention was to illustrate in a teaching device what has been discovered about the relationship between the different articulatory regions of the vocal tract: that laryngeal quality is fundamental (reflected in the vocal tract diagram and in the laryngoscopic videos and ultrasound images that accompany the Laryngeal consonants, and heard in voice qualities related to laryngeal constriction); that the control of articulatory and acoustic cues in infant speech acquisition originates in the larynx/pharynx (emulated by the emphasis placed on Voice Qualities in the app); and that the acquisition of the ability to produce consonants, let alone vowels, is inherently accompanied by coarticulation in the laryngeal area, i.e. in the varying of laryngeal states (illustrated in the app by the possibility of relating any of the qualities on the Voice Qualities page to vowel or consonant sets on the other pages) [2, 7]. This is not to claim that users of the app will follow the same process of phonetic learning that infants use in acquiring speech, any more than adult users are able to discriminate speech-sound categories perceptually to the same degree that infants can. However, offering learners the chance to practice qualities that emanate from different parts of the vocal tract and to relate them conceptually and compare them with vowel qualities or with consonant strictures that have the same articulatory origins can be defended as more natural than instruction in isolated consonants or vowels alone. Testing this hypothesis would require examining how the app complements classroom instruction (or indeed, private individual learning). An instructor could very well assist learners to interpret and compare auditory categories and visual images in the database, via the model, by switching between pages of Consonants, Vowels, and Voice Qualities and pointing out which sound elements co-occur and in what combinations in each particular speech system of the various languages of the world.

5. REFERENCES

[1] Coey, C., Esling, J. H., Moisik, S. R. 2014. *iPA Phonetics*, Version 1.0 [2014]. Department of Linguistics, University of Victoria.

[2] Benner, A., Grenon, I., Esling, J. H. 2007. Infants' phonetic acquisition of voice quality parameters in the first year of life. *Proc. 16th ICPhS* Saarbrücken, 2073–2076.

[3] Edmondson, J. A., Esling, J. H. 2006. The valves of the throat and their functioning in tone, vocal register, and stress: Laryngoscopic case studies. *Phonology* 23, 157–191.

[4] Edmondson, J. A., Padayodi, C. M., Hassan, Z. M., Esling, J. H. 2007. The laryngeal articulator: Source and resonator. *Proc. 16th ICPhS* Saarbrücken, 2065–2068.

[5] Esling, J. H. 2005. There are no back vowels: The Laryngeal Articulator Model. *Canadian Journal of Linguistics* 50, 13–44.

[6] Esling, J. H. 2010. Phonetic notation. In: Hardcastle, W. J., Laver, J., Gibbon, F. E. (eds), *The Handbook of Phonetic Sciences,* 2nd ed. Oxford: Wiley-Blackwell, 678–702.

[7] Esling, J. H. 2012. The articulatory function of the larynx and the origins of speech. *Proc. Annual Meeting Berkeley Linguistics Society* 38, University of California, Berkeley. (via *eLanguage* of the LSA)

[8] IPA. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

[9] Martínez-Celdrán, E. 2004. Problems in the classification of approximants. *JIPA* 34, 201–210.

[10] Miller, A. L., Brugman, J., Sands, B., Namaseb, L., Exter, M., Collins, C. 2009. Differences in airstream and posterior place of articulation among N|uu clicks. *JIPA* 39, 129–161.

[11] Moisik, S. R. 2013. *The Epilarynx in Speech*. PhD dissertation, University of Victoria.

[12] *Seeing Speech*. 2013. Carnegie Trust Ultrasound Tongue Imaging Resource, University of Glasgow and Queen Margaret, Strathclyde, Edinburgh, Aberdeen Universities. http://www.seeingspeech.arts.gla.ac.uk