



# Visible Vowels: a Tool for the Visualization of Vowel Variation

Wilbert Heeringa<sup>1</sup>, Hans Van de Velde<sup>1</sup>

<sup>1</sup>Fryske Akademy, The Netherlands

WHeeringa@fryske-akademy.nl, HVandeVelde@fryske-akademy.nl

## Abstract

This paper presents Visible Vowels, a web app that visualizes variation in  $f_0$ , formants and duration. It combines user friendliness with maximum functionality and flexibility, using a live plot view.

**Index Terms:**  $f_0$ , formants, duration, normalization, visualization

## 1. Introduction

Regional and sociophonetic vowel variation is frequently studied on the basis of vowel formants, which are the resonant frequencies of the vocal tract when a vowel is pronounced.

In sociophonetics, PLOTNIK and NORM are frequently used programs. PLOTNIK was developed by Labov. [1] It plots the vowels in F1/F2 space. Any selection of vowels can be visualized. Subsets of vowels can be hilited within the larger set. Means, standard deviations and the significance of difference between means can be calculated. Vowel systems of different speakers can be compared, by overlaying their vowel plots. PLOTNIK runs only on Macintosh operating systems.

The vowel normalization and plotting suite NORM is a web-based interface to the R package `vowels`. This package includes “procedures for the manipulation, normalization, and plotting of phonetic and sociophonetic vowel formant data.” [2]. The NORM web application especially offers the possibility to run a number of normalization techniques for normalizing formant data and to quickly compare the results. The web application NORM runs on any platform. The use of NORM does not require any knowledge of the programming language R, but is less flexible than the `vowels` package. Therefore, the authors encourage users to try the `vowels` package rather than using NORM.

Another tool for acoustic phonetic, sociolinguistic, and experimental phonological research is VOIS3D. The program offers the possibility to normalize formant frequencies and duration. Additionally, it provides an analytic geometric solution for assessment of spectral overlap. Normalized scatter for two vowel distributions is modeled as two best-fit ellipses oriented at angles with respect to the F1 and F2 axes. The output of the metric is an overlap fraction, which represents the region shared by both best-fit ellipses. This tool runs only on Windows operating systems. [3]

An important limitation of both PLOTNIK and NORM is that they only offer visualization of formants, not of duration and  $f_0$  of vowels. VOIS3D, however, can visualize vowel duration variation.

A package related to the `vowels` package is the `phonR` package developed by McCloy [4], offering the visualization of trajectories with an unlimited number of measure points. Additionally, IPA glyphs, ellipses showing degree of confidence in the location of the mean of each vowel/group, and convex hulls representing the outline of the vowel space can be added

to the vowel plot. Using this package it is possible to calculate and plot the degree of encroachment or overlap between vowel categories by means of a heat map.

In this paper we present Visible Vowels. Visible Vowels is a web application that aims to combine user friendliness with maximum flexibility and functionality. The web app uses McCloy’s [4] `phonR` package for the visualization of F1/F2 plots, but also includes modules for the visualization of  $f_0$  dynamics and duration. Different from earlier vowel plot programs Visible Vowels uses a live view, i.e. each time the user changes something in the settings, the plot shown in the viewer is immediately adjusted accordingly. This makes the comparison of, for example, different normalization techniques extremely easy.

## 2. Program

The web app Visible Vowels is implemented in the programming language R. [5]. The app uses the R packages: `phonR`, `gplots`, `plotrix`, `lattice`, `readxl`, `WriteXLS`, `DT`, `psych` and `pracma`. The user interface has been built using the packages `shiny` [6] and `shinyBS` [7]. Shiny is a web application framework for R and was developed by RStudio.

Visible Vowels is freely available at <https://fryske-akademy.nl/fa-apps/VisVow/> and works well in Firefox and Safari. In other browsers the program has not been tested so far. The app consists of five tab panels: ‘Load file’, ‘ $f_0$ ’, ‘Formants’, ‘Duration’ and ‘Help’. We discuss each of the panels below.

In the panels ‘ $f_0$ ’, ‘Formants’ and ‘Duration’ the user can download the (modified) input table either as tab-delimited text file or as Office Open XML file. Graphs that are generated in these panels can be saved in five different formats (SVG, PDF, EPS, JPG, PNG) and in three different sizes (small, medium, large).

### 2.1. Load file

In this panel, a spreadsheet file that contains the measurements is loaded. Spreadsheets are usually made in Microsoft Excel, therefore a file with Excel’s default file extension (‘.xlsx’) can be used without the need to export this file to a text file. In the spreadsheet, the first column should contain the speaker labels. An arbitrary number of columns which represent categorical variables such as location, language, gender, age group, etc. may follow. The last and mandatory column is ‘vowel’, which contains the vowel labels. Labels can be the IPA symbols in Unicode or key words that identify lexical sets [8].

The first column containing measurements is ‘duration’ and contains the duration of the vowels.

Finally, a set of five columns should follow: ‘time’, ‘ $f_0$ ’, ‘F1’, ‘F2’ and ‘F3’. The variable ‘time’ gives the time point within the vowel interval in seconds or milliseconds, i.e. it is assumed that the vowel interval starts at 0 (milli)seconds.  $f_0$ , F1, F2 and F3 should be measured at the time given in the col-

umn ‘time’. The program assumes that they are measured in Hertz and not normalized.

The set of five columns may be repeated as *many times* as the user wishes, but should occur at least one time. A set should always consist of five columns. If F3 measurements are not available, the cells of the ‘F3’ columns may be kept empty. If just one set of five columns is given, the column ‘time’ may be kept empty as well.

An example data table is provided in the ‘Help’ panel.

## 2.2. f0

In this panel the f0 contour can be visualized per vowel or per value of a categorical variable. Per vowel or per category the f0 contour is shown for the times given in the ‘time’ columns of the five-column sets. It is also possible to generate multiple line graphs, where the lines represent the categories of a second categorical variable. For example, when visualizing contour plots per vowel, we might want to see the contours of different regions in each plot.

Hertz values can be converted to Bark, ERB, Log, Mel and ST values.

## 2.3. Formants

In this panel vowels are plotted in F1/F2 space. The different colors of the plot symbols represent the categories of one categorical variable, and the shapes represent the categories of another categorical variable. Shapes can be circles, triangles, squares etc, but when ‘vowel’ is chosen, the vowel labels are shown instead.

It is possible to average categories of the categorical variables. For example, for 10 vowels, three regions and two genders the plot will show  $10 \times 3 \times 2 = 60$  plot symbols in three different colors and two different shapes. Additionally, for each categorical variable it is possible to select a subset of categories, with or without averaging the chosen category.

Ellipses showing the degree of confidence in the location of the mean of each vowel/group can be shown for different confidence levels ( $\pm 1, 1.96, 2, 2.58$  and 3 standard deviations). Convex hulls can be drawn in order to delineate the outline of the vowel space per category of a categorical variable.

When multiple five-column sets are included in the input table, vowel trajectories can be drawn. Given  $n$  sets, the user can include all time points or any subset of points, with a minimum of two subsets.

Hertz values can be converted to Bark, ERB, Log and Mel. There are 12 vowel normalization methods available: Gerstman’s range normalization, Lobanov’s  $z$ -score transformation, Labov’s ANAE methods 1 and 2, Miller’s formant ratio normalization, Nearey’s individual log-mean model, Nearey’s shared log-mean model, Syrdal & Gopal’s distance transformation, Tyler & Kendall’s difference metric, Watt & Fabricius’ S-centroid procedure, Watt & Fabricius’ modified S-centroid procedure, Convex hull centroid procedure. The latter method is similar to the Watt & Fabricius’ S-centroid procedure, but uses all the points the convex hull is made up of. Some methods require f0 and/or F3 measurements.

The normalization methods can be used on the basis of any scale (Hertz, Bark, etc.) except for the methods of Labov, Miller and Nearey, which can only be applied to Hertz values in our program. We decided thus since these methods use the log function in some steps of the normalization procedure.

Especially when a convex hull is drawn around the vowel space, the effects of the different normalization methods can

easily be compared due to the live plot view.

Most normalization methods calculate a scaling factor per speaker and per formant by which the original Hertz values are multiplied. The scaling factor can be calculated by considering data for all time points or any subset of time points. Thus it is possible to visualize formants at the 50% point which are normalized on the basis of data for the 20% and 80% point.

## 2.4. Duration

In the duration panel durations can be normalized by means of Lobanov’s  $z$ -score transformation. [9]

Average vowel durations for the categories of a categorical variable are shown by means of either a line plot or a bar graph. In the line graphs 95% confidence interval bars are included.

Using either multiple line graphs or grouped bar graphs it is possible to visualize a second categorical variable, where each line in a line graph represents a category. In a grouped bar graph, per category of the first categorical variable, bars are shown for all categories of the second categorical variable. Multiple line graphs do not show confidence interval bars.

## 2.5. Help

The help function provides an ‘about’ section, describes the format of the input file, gives information about the normalization methods and gives references for the R packages that are used in the program.

## 3. Future work

We will extend the program with the possibility to visualize F3 variation and vowel inherent spectral change (VISC). We will also add three-dimensional graphs in which, for example, F1 and F2 are combined with F3, VISC or duration. As in PLOTNIK we want to add significance tests for comparing means, and a measure that reflects the quality of vowel normalization in relation to the given data. Finally, we want to develop a stand-alone version of the program. In this version we will implement clickable maps, in which a user can click on a vowel in the F1/F2 space and hear the recording.

## 4. References

- [1] W. Labov, *A PLOTNIK tutorial*, 2002.
- [2] T. Kendall and E. R. Thomas, “Package ‘vowels’: Vowel Manipulation, Normalization, and Plotting,” 2015, version 1.2-1.
- [3] A. Wassink, “A geometric representation of spectral and temporal vowel features: Quantification of vowel overlap in three varieties,” *Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 2334–2350, 2006.
- [4] D. R. McCloy, “Package ‘phonR’: Tools for Phoneticians and Phonologists,” 2016, version 1.0-7.
- [5] R Core Team, “R: A language and environment for statistical computing,” Vienna, Austria, 2017, URL: <http://www.R-project.org/>.
- [6] W. Chang, J. Cheng, J. Allaire, Y. Xie, and J. McPherson, “Package ‘shiny’: Web Application Framework for R,” 2017, version 1.0.1.
- [7] E. Bailey, “Package ‘shinyBS’: Twitter Bootstrap Components for Shiny,” 2015, version 0.61.
- [8] J. C. Wells, *Accents of English*. Cambridge: Cambridge University Press, 1982, vol. 116.
- [9] H. Wang and V. J. van Heuven, “Acoustical analysis of English vowels produced by Chinese, Dutch and American speakers,” in *Linguistics in the Netherlands 2006*, J. van de Weijer and B. Los, Eds. Amsterdam: John Benjamins, 2006, pp. 237–248.