Relationships Between Prosodic-Linguistic Features and High-Level Descriptors of Speed Dates

Sidd Jagadish, Ranjay Krishna‡, Gabriele Carotti-Sha∓*

siddj@stanford.edu, rak248@stanford.edu, gcarotti@stanford.edu
∗Department of Statistics, Stanford University
‡Department of Computer Science, Stanford University
∓ Symbolic Systems, Stanford University

Abstract

We extract lexical and prosodic features of speech from 1493 speed dates at Stanford University. Each speed date is accompanied by information about what each participant thought of the other, including 1-10 assessments of courteousness and funniness. Here, we investigate prediction of these assessments from lexical and prosodic features. We find that for many labels, lexical and prosodic features can powerfully predict our high-level descriptors.

Index Terms: Prosody, Speed Date, AdaBoost, Gender differences in prosody, factor analysis, courteous, funny

1. Introduction

2. Related Works

3. Data

We were given, courtesy of Jurafsky (2013), a dataset consisting of 1980 (CHECK THIS VALUE) heterosexual speed dates. For each date, we have two .wav files corresponding to the microphones attached to each participant. We also have high-level descriptors of each participant, including their height, weight, and ethnicity. In addition to this, for each participant, we have 1-10 assessments of various qualities of both themselves and each other person they went on a speeddate with, including funniness and courteousness.

3.1. Feature Extraction

3.1.1. Prosodic Feature

Prosodic features were extracted using openSMILE. Lexical features were extracted in Python, with the help of the LIWC dictionary.

3.1.2. Lexical Feature

3.1.3. LIWC Feature

3.1.4. Accomodation Feature

Accommodation features include rate of

4. Exploratory Analysis

Before attempting to classify speakers as funny or not, we examine the underlying dimensions along which the data varies. To do, so we use exploratory factor analysis. To determine the number of factors, we use a scree test and various non-graphical measures, including parallel analysis and an optimal coordinates test, as described in (CITE). We find an optimal number of factors $k = 5$ for both males and females, conducted separately. Figure 1 shows the two scree plots.

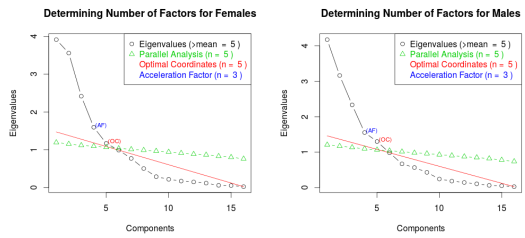

Figure 1: *Determining Number of Factors*

Interestingly, although we find that the various factors for male and female speech are similar, they explain different amounts of variation in the data. For males, the first factor reflects high intensity values and low intensity variation, explaining

5. Methodology

6. Classification Results

7. Analysis

8. Page layout and style

8.1. Figures

All figures must be centered on the column (or page, if the figure spans both columns). Figure captions should follow each figure and have the format given in Figure 2.

Figures should preferably be line drawings. If they contain gray levels or colors, they should be checked to print well on a high-quality non-color laser printer.

Graphics (i.e. illustrations, figures) must not use stipple fill patterns because they will not reproduce properly in Acrobat PDF. Please use only SOLID FILL COLORS.

Figures which span 2 columns (i.e. occupy full page width) should be placed at the top or bottom of the page.

8.2. Tables

An example of a table is shown as Table 1. Somewhat different styles are allowed according to the type and purpose of the table.

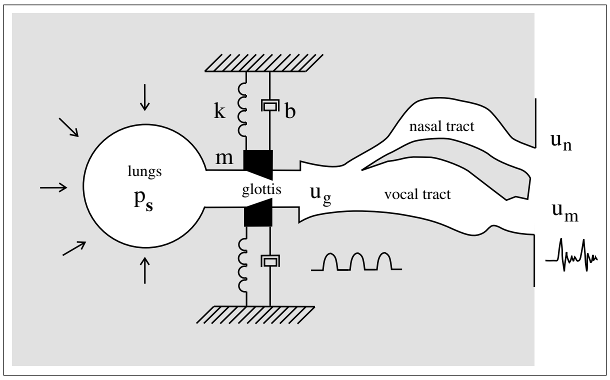

Figure 2: *Schematic diagram of speech production.*

The caption text may be above or below the table.

Table 1: *This is an example of a table.*

ratio	decibels
1/1	0
2/1	≈ 6
3.16	10
10/1	20
1/10	-20
100/1	40
1000/1	60

8.3. Equations

Equations should be placed on separate lines and numbered. Examples of equations are given below. Particularly,

$$x(t) = s(f_\omega(t)) \tag{1}$$

where $f_\omega(t)$ is a special warping function

$$f_\omega(t) = \frac{1}{2\pi j} \oint_C \frac{\nu^{-1k} d\nu}{(1 - \beta\nu^{-1})(\nu^{-1} - \beta)} \tag{2}$$

A residue theorem states that

$$\oint_C F(z)dz = 2\pi j \sum_k Res[F(z), p_k] \tag{3}$$

Applying (3) to (1), it is straightforward to see that

$$1 + 1 = \pi \tag{4}$$

Finally we have proven the secret theorem of all speech sciences. No more math is needed to show how useful the result is!

8.4. References

The reference format is the standard IEEE one. References should be numbered in order of appearance, for example [1], [2], and [3].

9. Future Work

10. Conclusions

11. Acknowledgements

We thank Dan Jurafsky for giving us access to the Speed Dating Dataset and for his continued guidance during this project.

12. References

[1] Smith, J. O. and Abel, J. S., “Bark and ERB Bilinear Transforms”, IEEE Trans. Speech and Audio Proc., 7(6):697–708, 1999.

[2] Soquet, A., Saerens, M. and Jospa, P.,“Acoustic-articulatory inversion”, in T. Kohonen [Ed], Artificial Neural Networks, 371-376, Elsevier, 1991.

[3] Stone, H.S., “On the uniqueness of the convolution theorem for the Fourier transform”, NEC Labs. Amer. Princeton, NJ. Online: http://citeseer.ist.psu.edu/176038.html, accessed on 19 Mar 2008.