Lecture 16: Parametric TTS, Intoxication, Depression, Trauma, Personality
Evaluation of TTS

- Intelligibility Tests
  - Diagnostic Rhyme Test (DRT)
    - Humans do listening identification choice between two words differing by a single phonetic feature
      - Voicing, nasality, sustenation, sibilation
      - 96 rhyming pairs
      - Veal/feel, meat/beat, vee/bee, zee/thee, etc
      - Subject hears “veal”, chooses either “veal” or “feel”
      - Subject also hears “feel”, chooses either “veal” or “feel”
    - % of right answers is intelligibility score.

- Overall Quality Tests
  - Have listeners rate space on a scale from 1 (bad) to 5 (excellent) (Mean Opinion Score)
  - AB Tests (prefer A, prefer B) (preference tests)

Huang, Acero, Hon
Parametric Synthesis

- Developed by Tokuda and Zen
- Proposed in mid-'90s, popular since 2007ish
- Big idea: Use classifiers/regressors to predict all of F0, duration, spectral envelope. Synthesize everything
- Initial work uses the same HMM we used for ASR, but in reverse
Parametric Synthesis

+ Small footprint
+ Don’t need huge amount of data to train
+ Flexible: easier to modify pitch for emotional change, or use MLLR adaptation to change voice characteristics
+ Smooth: no discontinuities in spectrum and prosody due to join artifacts
  - Too smooth: flat, monotone, spectral smearing in time
  - Vocoding effects: buzzy unnatural sound
HMM synthesis

\[ a_{ij} : \text{state transition probability} \]
\[ b_i(o_t) : \text{output probability} \]

Tokuda et al 2009
HTS system overview

(Tokuda, Zen, & Black. 2009)
What does the HMM produce?

Figure 10: ML-based excitation scheme proposed by Maia et al. for HMM-based speech synthesis: filters $H_v(z)$ and $H_u(z)$ are associated with each state.

(Tokuda, Zen, & Black. 2009)
Synthesis with source-filter model

Piece-wise stationary, source-filter generative model $p(x | o)$

Vocal source
- Pulse train (voiced)
- Fundamental frequency
- Aperiodicity voicing
- White noise (unvoiced)

Vocal tract filter
- Cepstrum, LPC, ...
- Overlap/shift windowing

Speech

$x(n) = h(n) * e(n)$
Key Questions in Parametric Synthesis

+ *What parameters do we predict?* Usually MFCCs for spectrum, log F0, voicing/excitation

+ *How do we combine them?* Exact parameterization and combining them well reduces robotic buzzy effects

+ *How do we make predictions?* Choice of HMM, machine learning approaches. Less important than the vocoding/combination issues
HTS Example

- Listen to the “low level” buzzy quality characteristic of most parametric systems
- Listen to clarity/impact of plosives compared to concatenative example
Comparing vocoder/excitation models

Figure 11: Waveforms from top to bottom: natural speech and its residual, speech and excitation synthesized with simple periodic pulse-train or white-noise excitation, speech and excitation synthesized with STRAIGHT vocoding method, and speech and excitation synthesized with ML excitation method.

(Tokuda, Zen, & Black. 2009)
End to end neural net synthesis

- TTS as a language model of individual samples

\[ p(x) = \prod_{t=1}^{T} p(x_t | x_1, \ldots, x_{t-1}) \]

- Condition on phoneme sequence / prosodic features \( h \)

\[ p(x | h) = \prod_{t=1}^{T} p(x_t | x_1, \ldots, x_{t-1}, h) \]

(Van den Oord et al. 2016)
Wavenet end to end synthesis

- TTS as a language model of individual samples

\[
p(x) = \prod_{t=1}^{T} p(x_t | x_1, \ldots, x_{t-1})
\]

- Condition on phoneme sequence / prosodic features \( h \)

\[
p(x | h) = \prod_{t=1}^{T} p(x_t | x_1, \ldots, x_{t-1}, h)
\]

(Van den Oord et al. 2016)
Causal convolution architecture

(Van den Oord et al. 2016)
Dilated causal convolutions

(Van den Oord et al. 2016)
Output encoding

- Real values for samples don’t work well
- Many quantizations already exist for speech (from telephony mostly)
- Output is a softmax classifier over 256 quantized values (mu law)

(Van den Oord et al. 2016)
Mean opinion score results

(Van den Oord et al. 2016)
Sequence to sequence with attention

- Predict frames and use a more standard vocoder
- Input is character sequences rather than phonetic features

(Wang et al. 2017)
Intoxication
Methods:

- 35 young adults, 19 males, 16 females
- given series of doses of alcohol
- speech collected at 4 BAC stages
  - Rainbow passage
  - difficult words (buttercup, shapupie)
  - extemp speech ("Tell us about your favorite TV program")
  - head-mounted mikes

Investigated:

- F0 mean and variance
- duration/rate of speech
- intensity
- disfluencies
Hollien et al 2001 Results:
F0
Hollien et al 2001 Results: Duration

<table>
<thead>
<tr>
<th>Group</th>
<th>Level of intoxication (BrAC)</th>
<th>Shift (0.00–0.12)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.00</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (s)</td>
<td>25.3</td>
<td>25.8</td>
</tr>
<tr>
<td>S.D. (s)</td>
<td>2.9</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (s)</td>
<td>25.1</td>
<td>25.5</td>
</tr>
<tr>
<td>S.D. (s)</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>Subjects</td>
<td>N</td>
<td>0.00</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Males</td>
<td>19</td>
<td>3.2</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>2.0</td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Females</td>
<td>16</td>
<td>2.2</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Mean</td>
<td>35</td>
<td>2.7</td>
</tr>
</tbody>
</table>
Hollien et al
2001 Results:
Magnitudes
Hollien et al 2001 Results: Speaker Specific Effects

- 20% of speakers did not follow these trends
A famous case study

The *Exxon Valdez* oil spill occurred in Prince William Sound, Alaska, on March 24, 1989, when the *Exxon Valdez*, an oil tanker bound for Long Beach, California, struck Prince William Sound's Bligh Reef and spilled 260,000 to 750,000 barrels (41,000 to 119,000 m³) of crude oil.[1][2] It is considered to be one of the most devastating human-caused environmental disasters.[3] As
Was Captain Hazelwood drunk?

- Not clear if this is relevant, since seems like other questionable corporate things were going on:
  - he was asleep below deck
  - The third mate was in charge of the wheelhouse
  - the ship’s radar was broken
- But is a well-studied case
Johnson et al examined 3 kinds of cues

- Segmental Effects (phoneme, syllable, word level)
- Disfluencies
- Suprasegmental Effects (stress, intonation, etc.)
Keith Johnsons /s/ and /ʃ/

Fig. 1. Power spectra of /s/ (a) and /ʃ/ (b) produced by K. J. in a quiet recording booth with recording equipment responsive up to 5,000 Hz.

e.g. “sun” vs “shun”
/ʃ/: Captain Hazelwood

"she’s"

"shout"

Fig. 2. Power spectra of /ʃ/ produced by Captain Hazelwood in the words she’s and shout recorded 33 h before the accident. Each spectrum is paired with a spectrum of the background noise from a nearby open-mike pause.
33 Hrs before

1 Hr before

Immediately after

1 Hr after

9 Hrs after
Fig. 2. Power spectra of /ʃ/ produced by Captain Hazelwood in the words she's and shout recorded 33 h before the accident. Each spectrum is paired with a spectrum of the background noise from a nearby open-mike pause.
Duration

Segment Durations of "Exxon Valdez"

- 33 Hrs before
- 1 Hr before
- Immediately after
- 1 Hr after
- 9 Hrs after

Duration (ms)
Table 3. Summary of phenomena found in the analysis of the NTSB tape (numbers in parentheses indicate the time of recording)

<table>
<thead>
<tr>
<th>Gross effects</th>
<th>revisions</th>
</tr>
</thead>
<tbody>
<tr>
<td>(−1) Exxon Ba, uh Exxon Valdez</td>
<td></td>
</tr>
<tr>
<td>(−1) departed disembarked</td>
<td></td>
</tr>
<tr>
<td>(−1) I, we’ll</td>
<td></td>
</tr>
<tr>
<td>(−1) columbia gla, columbia bay</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Segmental effects</th>
<th>misarticulation of /r/ and /l/</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0) northerly, little, drizzle, visibility</td>
<td></td>
</tr>
<tr>
<td>(/s/ becomes /ʃ/ (fig. 3)</td>
<td></td>
</tr>
<tr>
<td>final devoicing (e.g. /z/ → /s/)</td>
<td></td>
</tr>
<tr>
<td>(−1,0,+1) Valdez → Valdes</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Suprasegmental effects</th>
<th>reduced speaking rate (fig. 4, 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean change in pitch range (talker-dependent, fig. 6)</td>
<td></td>
</tr>
<tr>
<td>increased F₀ jitter (fig. 6)</td>
<td></td>
</tr>
</tbody>
</table>
Problems

- If intoxicated speech, why wasn’t s pronounced as sh 1 hour before?
- Other kinds of speaker state could cause drop in F0, slower speech, and disfluencies?
  - Stress, just having woken up, trauma....
Automatic Classification

- Use of prosodic speech characteristics for automated detection of alcohol intoxication
  - Michael Levit, Richard Huber, Anton Batliner, Elmar Noeth
- Break utterance into phrases automatically, based on
  - fundamental frequency (where possible);
  - zero-crossing rate
Then use 4 classes of features

- Prosodic
  - F0 max, F0 min, energy max, energy min, pause length
- Duration of voiced regions, unvoiced regions, etc.
- Jitter and shimmer
  - Jitter is variation in pitch
  - Shimmer is variation in energy
- Average cepstrum and cepstral slope
Methods

- Alcoholized speech samples collected at the Police Academy of Hessen, Germany
- 120 readings (87 minutes) of a fable
- 33 male speakers
- BAC between 0 and .24/mille

<table>
<thead>
<tr>
<th>Alcohol Blood Level</th>
<th>0.0</th>
<th>&lt; 0.4</th>
<th>&lt; 0.8</th>
<th>&lt; 1.2</th>
<th>&lt; 1.6</th>
<th>&lt; 2.0</th>
<th>&lt; 2.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recordings</td>
<td>32</td>
<td>20</td>
<td>20</td>
<td>18</td>
<td>20</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

- Binary task: above or below 0.8/mille
- leave-one-out cross-validation
- neural net classifier
Results of Levit et al.

- Used dev set to find best classifier
- This suggested two feature classes:
  - Prosodic features
  - Jitter/shimmer
- Results with this classifier
  - 62% phrase-accuracy
  - 69% for the whole speech sample
  - voting of the phrases
New Corpus!

- Alcohol Language Corpus
  - Florian Schiel et al 2009, 2010
  - [http://www.bas.uni-muenchen.de/forschung/Bas/BasALCeng.html](http://www.bas.uni-muenchen.de/forschung/Bas/BasALCeng.html)
- 162 speakers (77 female, 85 male)
  - recorded in a car (sometimes with engine running)
  - command and control speech (“turn off the radio”)
  - spontaneous dialogue, monologue, question answering
  - read speech
  - counts of disfluencies, etc

- sample, drunk:
- sample, sober:
Automatic detection in ALC: Paralinguistic Challenge 2011

- Human: 66-72% (Schiel 2011, Ultes, Schmitt, Minker 2011)
- Machine: roughly 65%-70%
- Example features from winning system:


- Prosody (f0, duration, energy, jitter, shimmer)
- Spectral (MFCC, MFB log-energy, formants)
- Computed over whole utterance and small windows
- normalized phoneme duration
- iterative speaker normalization
Depression
Stirman and Pennebaker

- Suicidal poets
- 300 poems from early, middle, late periods of
  - 9 suicidal poets
  - 9 non-suicidal poets
Stirman and Pennebaker: 2 models

- **Durkheim disengagement model:**
  - suicidal individual has failed to integrate into society sufficiently, is detached from social life
  - detach from the source of their pain, withdraw from social relationships, become more self-oriented
  - prediction:
    - more self-reference, less group references

- **Hopelessness model:**
  - Suicide takes place during extended periods of sadness and desperation, pervasive feelings of helplessness, thoughts of death
  - prediction:
    - more negative emotion, fewer positive, more refs to death
Methods

• 156 poems from 9 poets who
  • committed suicide
  • published, well-known
  • in English
  • have written within 1 year of committing suicide
• Control poets matched for nationality, education, sex, era.
### TABLE 2. Suicidal Poets and Their Controls

<table>
<thead>
<tr>
<th>Suicidal Poet</th>
<th>Age at Death</th>
<th>Control Poet</th>
<th>Cutoff Age</th>
<th>Nationality</th>
<th>Other Similarities</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Berryman (1914–1972)</td>
<td>58</td>
<td>Lawrence Ferlinghetti (1919–)</td>
<td>59 (1978)</td>
<td>American</td>
<td>PhD</td>
</tr>
<tr>
<td>Adam L. Gordon (1833–1870)</td>
<td>37</td>
<td>Matthew Arnold (1822–1888)</td>
<td>45 (1867)</td>
<td>British</td>
<td></td>
</tr>
<tr>
<td>Sarah Teasdale (1884–1933)</td>
<td>49</td>
<td>Edna St. V. Millay (1892–1950)</td>
<td>49 (1941)</td>
<td>American</td>
<td></td>
</tr>
<tr>
<td>Hart Crane (1899–1932)</td>
<td>33</td>
<td>Joyce Kilmer (1886–1918)</td>
<td>32 (1918)</td>
<td>American</td>
<td></td>
</tr>
<tr>
<td>Sergei Esenin (1895–1925)</td>
<td>30</td>
<td>Boris Pasternak (1890–1960)</td>
<td>35 (1930)</td>
<td>Russian</td>
<td></td>
</tr>
<tr>
<td>Vladimir Mayakovsky (1893–1930)</td>
<td>37</td>
<td>Osip Mandelstam (1891–1938)</td>
<td>37 (1928)</td>
<td>Russian</td>
<td></td>
</tr>
</tbody>
</table>


TABLE 1. Means for LIWC Categories

<table>
<thead>
<tr>
<th></th>
<th>Suicide Group</th>
<th></th>
<th>Control Group</th>
<th></th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Early</td>
<td>Middle</td>
<td>Later</td>
<td>Early</td>
<td>Middle</td>
</tr>
<tr>
<td>Disengagement theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I (me, my)</td>
<td>4.0</td>
<td>3.4</td>
<td>4.0</td>
<td>2.5</td>
<td>1.6</td>
</tr>
<tr>
<td>We (us, our)</td>
<td>.73</td>
<td>1.3</td>
<td>.85</td>
<td>.69</td>
<td>.40</td>
</tr>
<tr>
<td>Communication (talk, share)</td>
<td>1.2</td>
<td>1.1</td>
<td>1.0</td>
<td>.89</td>
<td>1.1</td>
</tr>
<tr>
<td>Hopelessness theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative emotion (hate, worthless)</td>
<td>2.2</td>
<td>1.8</td>
<td>1.7</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>Positive emotion (happy, love)</td>
<td>3.3</td>
<td>3.1</td>
<td>3.9</td>
<td>2.9</td>
<td>2.9</td>
</tr>
<tr>
<td>Death (dead, grave)</td>
<td>.52</td>
<td>.47</td>
<td>.69</td>
<td>.34</td>
<td>.43</td>
</tr>
<tr>
<td>Other findings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual words (lust, breast)</td>
<td>.60</td>
<td>.84</td>
<td>.47</td>
<td>.36</td>
<td>.36</td>
</tr>
</tbody>
</table>

Note: Means reflect percentage of total words used in each poem within the relevant category. Effects refer to: S = suicide vs. nonsuicide main effect, P = phase of career main effect. All effects are significant $p \leq .05$, except ** $p \leq .08$. 
Significant factors

• Disengagement theory
  • I, me, mine
  • we, our, ours
• Hopelessness theory
  • death, grave
• Other
  • sexual words (lust, breast)
Beck (1967) cognitive theory of depression
- depression-prone individuals see the world and themselves in pervasively negative terms

Pyszynski and Greenberg (1987)
- think about themselves
- after the loss of a central source of self-worth, unable to exit a self-regulatory cycle concerned with efforts to regain what was lost.
- results in self-focus, self-blame

Durkheim social integration/disengagement
- perception of self as not integrated into society is key to suicidality and possibly depression
Methods

- College freshmen
  - 31 currently-depressed (standard inventories)
  - 26 formerly-depressed
  - 67 never-depressed
- Session 1: take depression inventory
- Session 2: write essay
  - please describe your deepest thoughts and feelings about being in college... write continuously off the top of your head. Don’t worry about grammar or spelling. Just write continuously.
Results

- depressed used more “I, me” than never-depressed
  - turned out to be only “I”
- and used more negative emotional words
- not enough “we” to check Durkheim model
- formerly depressed participants used more “I” in the last third of the essay
Ramirez-Esparza et al: Depression in English and Spanish

- Study 1: Use LIWC counts on posts from 320 English and Spanish forums
  - 80 posts each from depression forums in English and Spanish
  - 80 control posts each from breast cancer forums
- Run the following LIWC categories
  - I
  - we
  - negative emotion
  - positive emotion
## Results of Study 1

<table>
<thead>
<tr>
<th>Categories</th>
<th>English</th>
<th></th>
<th>Spanish</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=80</td>
<td>N=80</td>
<td>N=80</td>
<td>N=80</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>(SD)</td>
<td>(SD)</td>
<td>(SD)</td>
<td>(SD)</td>
</tr>
<tr>
<td>First person singular</td>
<td>12.24</td>
<td>4.03</td>
<td>9.30</td>
<td>5.03</td>
</tr>
<tr>
<td></td>
<td>(2.97)</td>
<td>(3.01)</td>
<td>(2.34)</td>
<td>(2.76)</td>
</tr>
<tr>
<td>First person plural</td>
<td>.18</td>
<td>.72</td>
<td>.22</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>(.33)</td>
<td>(1.06)</td>
<td>(.39)</td>
<td>(1.28)</td>
</tr>
<tr>
<td>Positive Emotions</td>
<td>1.72</td>
<td>2.54</td>
<td>2.99</td>
<td>3.53</td>
</tr>
<tr>
<td></td>
<td>(1.14)</td>
<td>(1.72)</td>
<td>(1.36)</td>
<td>(1.93)</td>
</tr>
<tr>
<td>t-value</td>
<td>-17.39*</td>
<td>-10.54*</td>
<td>-4.36*</td>
<td>5.32*</td>
</tr>
<tr>
<td>t-value</td>
<td>3.56*</td>
<td>2.04†</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Case Study: Online Forum Posts

- From depression forums:
  - 404 English posts
  - 404 Spanish posts
- Create a term by document matrix of content words
  - 200 most frequent content words
- Do a factor analysis
  - dimensionality reduction in term-document matrix
  - Used 5 factors
### English Factors

<table>
<thead>
<tr>
<th>FACTOR 1: Treatment</th>
<th>FACTOR 1: Disclosure</th>
<th>FACTOR 3: Family</th>
<th>FACTOR 4: Symptoms</th>
<th>FACTOR 5: School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication</td>
<td>Tell</td>
<td>Mom</td>
<td>Sleep</td>
<td>Constant</td>
</tr>
<tr>
<td>Effect</td>
<td>People</td>
<td>Daughter</td>
<td>Hour</td>
<td>Relationship</td>
</tr>
<tr>
<td>Depression</td>
<td>Know</td>
<td>Child</td>
<td>Food</td>
<td>School</td>
</tr>
<tr>
<td>Side</td>
<td>Happy</td>
<td>Family</td>
<td>Wake</td>
<td>High</td>
</tr>
<tr>
<td>Week</td>
<td>Talk</td>
<td>Brother</td>
<td>Morning</td>
<td>Lack</td>
</tr>
<tr>
<td>Therapy</td>
<td>Feel</td>
<td>Sister</td>
<td>Night</td>
<td>University</td>
</tr>
<tr>
<td>Suffer</td>
<td>Want</td>
<td>Dad</td>
<td>Bed</td>
<td>Social</td>
</tr>
<tr>
<td>Disorder</td>
<td>Suppose</td>
<td>Son</td>
<td>Stay</td>
<td>College</td>
</tr>
<tr>
<td>Doctor</td>
<td>Read</td>
<td>Love</td>
<td>Weight</td>
<td>Move</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>Hurt</td>
<td>Girl</td>
<td>Eat</td>
<td>Friend</td>
</tr>
<tr>
<td>Experience</td>
<td>Wrong</td>
<td>Young</td>
<td>Place</td>
<td>Girlfriend</td>
</tr>
<tr>
<td>Major</td>
<td>Emotional</td>
<td>Parent</td>
<td>Class</td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td>Mind</td>
<td>House</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatrist</td>
<td>Sad</td>
<td>Husband</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Make</td>
<td>Crazy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Spanish Factors

### FACTOR 1: Family
- MADRE/mother: .48
- HERMANO/brother: .46
- ABUELA/grandmother: .44
- PADRES/parents: .44
- PAPA/father: .43
- HORRIBLE/horrible: .39
- CASA/house: .39
- SUICIDIO/suicide: .36
- DINERO/money: .34
- ESTUDIOS/studies: .33
- CLASE/class: .32
- FEA/ugly: .32
- ASCO/disgust: .32
- COMER/eat: .31
- PEQUENIOO/small: .31
- FAMILIA/family: .31

### FACTOR 2: Relationship History
- RELACION/relationship: .51
- ENAMORADO/love: .48
- CONOCI/met: .45
- HABLAR/talk: .45
- CHICO/guy: .44
- AMIGOS/friends: .44
- NOVIO/boyfriend: .43
- JUNTOS/together: .37
- ESPECIAL/special: .36
- EMPECE/start: .33
- TIEMPO/time: .33
- DEJAR/leave: .33
- FINAL/end: .33
- HISTORIA/history: .31
- MESES/months: .31
- LLEGAR/arrive: .31

### FACTOR 3: Hopelessness
- NOCHE/night: .45
- SEGUNDO/second: .44
- MORIR/die: .42
- PAZ/peace: .39
- OJOS/eyes: .39
- ESPERO/hope: .38
- TERRIBLE/terrible: .36
- FUERTE/strong: .34
- CORAZON/heart: .31
- SUENIOS/dreams: .31

### FACTOR 4: School
- TIMIDA/timid: .41
- CARRERA/career: .41
- COLEGIO/school: .40
- CONFIANZA/trust: .37
- ESTUDIOS/studies: .37
- INCAPAZ/incapable: .33
- UNIVERSI./university: .33
- TONTERIA/foolishness: .32

### FACTOR 5: Treatment
- PSICOLOGO/psychologist
- ANSIEDAD/anxiety
- EMPRESA/company
- ANTIDEP/antidepressants
- SINTOMAS/symptoms
- MEDICAMENTO/medicines
Speech features for Depression


Commonly used features:

- F0 variance (monopitch)
- loudness variance (monoloudness)
- rate of speech (slower)
  - response delay, pauses
- spectral features
Social stage model of collective coping (Pennebaker & Harber, 1993). After a traumatic experience:

Stage 1: people cope by sharing their thoughts about the upsetting experience
Stage 2, a few weeks later: decrease in talking, but still thinking about event
Stage 3: 6-8 weeks later: reduction in both talking and thinking

What are the linguistic characteristics of stage 1?
Cohn, Mehl, Pennebaker: Linguistic Markers of Psychological Change Surrounding September 11, 2001

- 1084 LiveJournal users
- all blog entries for 2 months before and after 9/11
- Lumped prior two months into one “baseline” corpus.
- Investigated changes after 9/11 compared to that baseline
- Using LIWC categories
Variables examined

- Emotional positivity
  - difference between LIWC scores for positive emotion words (happy, good, nice) and negative emotion words (kill, ugly, guilty).

- Cognitive processing
  - think, question, because: concerned with organizing and intellectually understanding issues

- Social orientation
  - talk, share, friends and personal pronouns besides I/me. (essentially counts # of references to other people)
Last factor: Psychological Distancing

- psychological distancing
  - factor-analytic:
    - + articles,
    - + words > 6 letters long
    - - I/me/mine
    - - would/should/could
    - - present tense verbs
  - low score = personal, experiential lg, focus on here and now
  - high score: abstract, impersonal, rational tone
Results

A. Emotional Positivity

B. Cognitive Processes

C. Social Processes

D. Psychological Distancing
LiveJournal.com September 11, 2001 study: Positive and negative emotion words

Livejournal.com:

*l, me, my* on or after Sep 11, 2001


Graph from Pennebaker slides
September 11 LiveJournal.com study: We, us, our


Graph from Pennebaker slides
Trauma after Princess Diana’s death

- Princess Diana died August 30, 1997
- Over the next 4 weeks, scraped all conversations from “The UK Experience” chat room on AOL.
- 121 chat sessions among 3,139 participants.
- Compared to baseline rates:
  - Increase in *we*
  - Decrease in *I*
  - Increase in *negative emotional words*
Texas A&M Bonfire tragedy

- Gortner and Pennebaker
- Examined student newspaper in the weeks after the tragedy:
  - Increase in *we*
  - Increase in *I*
  - Increase in *negative emotion*
Another domain of trauma? Restaurant Reviews

Jurafsky, Chahuneau, Routledge, Smith 2014.

• 6562 restaurants
  • 900K reviews [www.yelp.com](http://www.yelp.com)

• Negative (★):
  • The bartender... absolutely horrible... we waited 10 min before we even got her attention... and then we had to wait 45 - FORTY FIVE! - minutes for our entrees... stalk the waitress to get the cheque... she didn't make eye contact or even break her stride to wait for a response...
What makes a bad review bad?

- **Negative sentiment language**
  - horrible awful terrible worst bad disgusting

- **narrative**
  - past tense
    - waited, didn’t make eye contact, was disappointing.

- **3rd person pronouns**
  - he she his her

- **other people**
  - manager, customer, minutes, money, waitress, waiter, bill, attitude, management, business, apology, mistake, table, charge, order, hostess,

- **mentions of we and us**

  we waited 10 min before we even got her attention... and then we had to wait 45 - FORTY FIVE! - minutes for our entrees... ...
We just saw texts with these characteristics!

- **Negative sentiment, past tense narratives about others**
- **Enormous increase in “we” and “us”: solace in community**
- **Chat group discussions after Princess Diana’s death**

- **Blog posts after September 11, 2001**

- **Student newspaper reports after a campus tragedy**

- **Conclusion:** *Awful reviews are trauma narratives*
Personality
Scherer’s typology of affective states

**Emotion**: relatively brief episode of synchronized response of all or most organismic subsystems in response to the evaluation of an external or internal event as being of major significance

   angry, sad, joyful, fearful, ashamed, proud, desperate

**Mood**: diffuse affect state ...change in subjective feeling, of low intensity but relatively long duration, often without apparent cause

   cheerful, gloomy, irritable, listless, depressed, buoyant

**Interpersonal stance**: affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange

   distant, cold, warm, supportive, contemptuous

**Attitudes**: relatively enduring, affectively colored beliefs, preferences predispositions towards objects or persons

   liking, loving, hating, valuing, desiring

**Personality traits**: emotionally laden, stable personality dispositions and behavior tendencies, typical for a person

   nervous, anxious, reckless, morose, hostile, envious, jealous
Personality and Cultural Values

- Personality refers to the structures and propensities inside a person that explain his or her characteristic patterns of thought, emotion, and behavior.
- Personality captures what people are like.
- Traits are defined as recurring regularities or trends in people’s responses to their environment.
- Cultural values, defined as shared beliefs about desirable end states or modes of conduct in a given culture, influence the expression of a person’s traits.
The Big Five Dimensions of Personality

- Extraversion vs. Introversion
  (sociable, assertive, playful vs. aloof, reserved, shy)
- Emotional stability vs. Neuroticism
  (calm, unemotional vs. insecure, anxious)
- Agreeableness vs. Disagreeable
  (friendly, cooperative vs. antagonistic, faultfinding)
- Conscientiousness vs. Unconscientious
  (self-disciplined, organised vs. inefficient, careless)
- Openness to experience
  (intellectual, insightful vs. shallow, unimaginative)
Aside: Do Animals Have Personalities?

- Gosling (1998) studied spotted hyenas. He:
  - had human observers use personality scales to rate the different hyenas in the group
  - did a factor analysis on these findings
  - found five dimensions
    three closely resembled the Big Five traits of neuroticism, openness to experience, and agreeableness
The Big Five Personality Traits

- Conscientiousness - dependable, organized, reliable, ambitious, hardworking, and persevering.
The Big Five Personality Traits, Cont’d

- Agreeableness - warm, kind, cooperative, sympathetic, helpful, and courteous.
- Strong desire to obtain acceptance in personal relationships as a means of expressing personality.
- Agreeable people focus on “getting along,” not necessarily “getting ahead.”
The Big Five Personality Traits, Cont’d

- Extraversion - talkative, sociable, passionate, assertive, bold, and dominant.
  - Easiest to judge in zero acquaintance situations — situations in which two people have only just met.
  - Prioritize desire to obtain power and influence within a social structure as a means of expressing personality.
  - High in positive affectivity — a tendency to experience pleasant, engaging moods such as enthusiasm, excitement, and elation.
The Big Five Personality Traits:

Neuroticism - nervous, moody, emotional, insecure, jealous.

- experience unpleasant moods such as hostility, nervousness, and annoyance.
- more likely to appraise day-to-day situations as stressful.
- less likely to believe they can cope with the stressors that they experience.
- related to locus of control (attribute causes of events to themselves or to the external environment)
  - neurotics hold an external locus of control: believe that the events that occur around them are driven by luck, chance, or fate.
  - less neurotic people hold internal locus of control: believe that their own behavior dictates events.
# External and Internal Locus of Control

<table>
<thead>
<tr>
<th>People with an External Locus of Control Tend to Believe:</th>
<th>People with an Internal Locus of Control Tend to Believe:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Many of the unhappy things in people's lives are partly due to bad luck.</td>
<td>People's misfortunes result from the mistakes they make.</td>
</tr>
<tr>
<td>Getting a good job depends mainly on being in the right place at the right time.</td>
<td>Becoming a success is a matter of hard work; luck has little or nothing to do with it.</td>
</tr>
<tr>
<td>Many times exam questions tend to be so unrelated to course work that studying is really useless.</td>
<td>In the case of the well-prepared student, there is rarely if ever such a thing as an unfair test.</td>
</tr>
<tr>
<td>This world is run by the few people in power, and there is not much the little guy can do about it.</td>
<td>The average citizen can have an influence in government decisions.</td>
</tr>
<tr>
<td>There's not much use in trying too hard to please people; if they like you, they like you.</td>
<td>People are lonely because they don't try to be friendly.</td>
</tr>
</tbody>
</table>
The Big Five Personality Traits, Cont’d

**Openness to experience** - curious, imaginative, creative, complex, refined, and sophisticated.

- Also called “Inquisitiveness” or “Intellectualness” or even “Culture.”
- high levels of creativity, the capacity to generate novel and useful ideas and solutions.
- Highly open individuals are more likely to migrate into artistic and scientific fields.
Changes in Big Five Dimensions Over the Life Span

McGraw-Hill/Irwin Chapter 9
Take the Big Five Inventory

http://www.outofservice.com/bigfive/
Corpora for studying personality:
Natural speech

**Electronically Activated Recorder (EAR)**


- a modified digital voice recorder that periodically records brief snippets of ambient sounds
- Attaches to the belt or in a purse-like bag while participants go about their daily lives.
Analog EAR-1: 90 minute tape
1997-2000
Digital EAR-2: digital voice recorder, flash drive 2001-2004
Mairesse et al. Two Corpora

- Pennebaker and King (1999)
  - 2,479 essays from psychology students (1.9 million words), “write whatever comes into your mind” for 20 minutes

- Mehl et al. (2006)
  - Speech from Electronically Activated Recorder (EAR)
  - Random snippets of conversation recorded, transcribed
  - 96 participants, total of 97,468 words and 15,269 utterances).
Mehl et al. (2006) data


- 96 psych freshman at UT Austin took the 44-item Big Five Inventory
- Agreed to wear EAR two weekdays continuously (when awake)
  - External mike clipped to collar
- 30-s on, 12.5-min off cycle = 4.8 recordings/hour
  - They were told they could erase anything they didn’t want researchers to hear
  - afterwards they reported wearing about 75% of their waking time
- Each sound file
  - transcribed
  - coded for environmental situation (location, activity)
  - 23 LIWC variables coded
  - 18 trained students listened to the files and assigned Big Five Inventory scores
### Ears (speech) corpus

<table>
<thead>
<tr>
<th>Introvert</th>
<th>Extravert</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Yeah you would do kilograms. Yeah I see what you’re saying.</td>
<td>- That’s my first yogurt experience here.</td>
</tr>
<tr>
<td>- On Tuesday I have class. I don’t know.</td>
<td>Really watery. Why?</td>
</tr>
<tr>
<td>- I don’t know. A16. Yeah, that is kind of cool.</td>
<td>- Damn. New game.</td>
</tr>
<tr>
<td>- I don’t know. I just can’t wait to be with you and not have to do this every night, you know?</td>
<td>- Oh.</td>
</tr>
<tr>
<td>- Yeah. You don’t know. Is there a bed in there? Well ok just...</td>
<td>- That’s so rude. That.</td>
</tr>
<tr>
<td></td>
<td>- Yeah, but he, they like each other.</td>
</tr>
<tr>
<td></td>
<td>He likes her.</td>
</tr>
<tr>
<td></td>
<td>- They are going to end up breaking up and he’s going to be like.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Unconscientious</th>
<th>Conscientious</th>
</tr>
</thead>
<tbody>
<tr>
<td>- With the Chinese. Get it together.</td>
<td>- I don’t, I don’t know for a fact but I would imagine that historically women who have entered prostitution have done so, not everyone, but for the majority out of extreme desperation and I think. I don’t know, i think people understand that desperation and they don’t see [...]</td>
</tr>
<tr>
<td>- I tried to yell at you through the window.</td>
<td></td>
</tr>
<tr>
<td>Oh. xxxx’s fucking a dumb ass. Look at him. Look at him, dude. Look at him. I wish we had a camera. He’s fucking brushing his t-shirt with a tooth brush. Get a kick of it. Don’t steal nothing.</td>
<td></td>
</tr>
</tbody>
</table>
# Essays corpus

<table>
<thead>
<tr>
<th>Introvert</th>
<th>Extravert</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’ve been waking up on time so far. What has it been, 5 days? Dear me, I’ll never keep it up, being such not a morning person and all. But maybe I’ll adjust, or not. I want internet access in my room, I don’t have it yet, but I will on Wed?? I think. But that ain’t soon enough, cause I got calculus homework [...]</td>
<td>I have some really random thoughts. I want the best things out of life. But I fear that I want too much! What if I fall flat on my face and don’t amount to anything. But I feel like I was born to do BIG things on this earth. But who knows… There is this Persian party today.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neurotic</th>
<th>Emotionally stable</th>
</tr>
</thead>
<tbody>
<tr>
<td>One of my friends just barged in, and I jumped in my seat. This is crazy. I should tell him not to do that again. I’m not that fastidious actually. But certain things annoy me. The things that would annoy me would actually annoy any normal human being, so I know I’m not a freak.</td>
<td>I should excel in this sport because I know how to push my body harder than anyone I know, no matter what the test I always push my body harder than everyone else. I want to be the best no matter what the sport or event. I should also be good at this because I love to ride my bike.</td>
</tr>
</tbody>
</table>
# Sample Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger words</td>
<td>LIWC</td>
<td>hate, kill, pissed</td>
</tr>
<tr>
<td>Metaphysical issues</td>
<td>LIWC</td>
<td>God, heaven, coffin</td>
</tr>
<tr>
<td>Physical state/function</td>
<td>LIWC</td>
<td>ache, breast, sleep</td>
</tr>
<tr>
<td>Inclusive words</td>
<td>LIWC</td>
<td>with, and, include</td>
</tr>
<tr>
<td>Social processes</td>
<td>LIWC</td>
<td>talk, us, friend</td>
</tr>
<tr>
<td>Family members</td>
<td>LIWC</td>
<td>mom, brother, cousin</td>
</tr>
<tr>
<td>Past tense verbs</td>
<td>LIWC</td>
<td>walked, were, had</td>
</tr>
<tr>
<td>References to friends</td>
<td>LIWC</td>
<td>pal, buddy, coworker</td>
</tr>
<tr>
<td>Imagery of words</td>
<td>MRC</td>
<td>Low: future, peace - High: table, car</td>
</tr>
<tr>
<td>Syllables per word</td>
<td>MRC</td>
<td>Low: a - High: uncompromisingly</td>
</tr>
<tr>
<td>Concreteness</td>
<td>MRC</td>
<td>Low: patience, candor - High: ship</td>
</tr>
<tr>
<td>Frequency of use</td>
<td>MRC</td>
<td>Low: duly, nudity - High: he, the</td>
</tr>
</tbody>
</table>
LIWC FEATURES (Pennebaker et al., 2001):

- **Standard counts:**
  - Word count (WC), words per sentence (WPS), type/token ratio (Unique), words captured (Dic), words longer than 6 letters (Sixltr), negations (Negate), assents (Assent), articles (Article), prepositions (Preps), numbers (Number)
  - Pronouns (Pronoun): 1st person singular (I), 1st person plural (We), total 1st person (Self), total 2nd person (You), total 3rd person (Other)

- **Psychological processes:**
  - Affective or emotional processes (Affect): positive emotions (Posemo), positive feelings (Posfeel), optimism and energy (Optim), negative emotions (Negemo), anxiety or fear (Anx), anger (Anger), sadness (Sad)
  - Cognitive Processes (Cogmech): causation (Cause), insight (Insight), discrepancy (Discrep), inhibition (Inhib), tentative (Tentat), certainty (Certain)
  - Sensory and perceptual processes (Senses): seeing (See), hearing (Hear), feeling (Feel)
  - Social processes (Social): communication (Comm), other references to people (Othref), friends (Friends), family (Family), humans (Humans)

- **Relativity:**
  - Time (Time), past tense verb (Past), present tense verb (Present), future tense verb (Future)
  - Space (Space): up (Up), down (Down), inclusive (Incl), exclusive (Excl)
  - Motion (Motion)

- **Personal concerns:**
  - Occupation (Occup): school (School), work and job (Job), achievement (Achieve)
  - Leisure activity (Leisure): home (Home), sports (Sports), television and movies (TV), music (Music)
  - Money and financial issues (Money)
  - Metaphysical issues (Metaph): religion (Relig), death (Death), physical states and functions (Posyscal), body states and symptoms (Body), sexuality (Sexual), eating and drinking (Eating), sleeping (Sleep), Grooming (Groom)

- **Other dimensions:**
  - Punctuation (Allpct): period (Period), comma (Comma), colon (Colon), semi-colon (Semic), question (Qmark), exclamation (Exclam), dash (Dash), quote (Quote), apostrophe (Apostro), parenthesis (Parenth), other (Otherp)
  - Swear words (Swear), nonfluencies (Nonfl), fillers (Fillers)
Utterance type

Labeled by parsing each utterance and then using heuristic rules based on parse tree:

**Commands**: imperatives, “can you”, etc.

**Backchannels**: yeah, ok, uh-huh, huh

**Questions**

**Assertions** (anything else)
Prosodic features

*Computed via Praat*

- **pitch** (mean, min, max, sd):
- **intensity** (mean, min, max, sd)
- **voiced time**
- **rate of speech** (words/second)
Classifiers from Weka

- **Classification (binary)**
  - C4.5 Decision Tree (J48)
  - Nearest neighbor
  - Naïve Bayes
  - Ripper
  - Adaboost
  - SVM with linear kernels

- **Regression (predict Likert values)**
  - linear regression
  - M5’ regression tree
  - SVMOreg

- **Ranking (training set T of ordered pairs)**
  - \( T = \{(x,y) | x, y, \text{are language samples from two individuals, } x \text{ has a higher score than } y \text{ for that personality trait}\} \)
  - Rankboost
### Ears (speech) corpus

<table>
<thead>
<tr>
<th>Data</th>
<th>Trait</th>
<th>Base</th>
<th>J48</th>
<th>NN</th>
<th>NB</th>
<th>JRIP</th>
<th>ADA</th>
<th>SMO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs</td>
<td>Extra</td>
<td>47.78</td>
<td>66.78</td>
<td>59.33</td>
<td>73.00●</td>
<td>60.44</td>
<td>73.00●</td>
<td>65.78</td>
</tr>
<tr>
<td>Obs</td>
<td>Emot</td>
<td>51.11</td>
<td>62.56</td>
<td>58.22</td>
<td>73.89●</td>
<td>56.22</td>
<td>48.78</td>
<td>60.33</td>
</tr>
<tr>
<td>Obs</td>
<td>Agree</td>
<td>47.78</td>
<td>48.78</td>
<td>51.89</td>
<td>61.33●</td>
<td>51.89</td>
<td>52.89</td>
<td>56.33</td>
</tr>
<tr>
<td>Obs</td>
<td>Conscl</td>
<td>47.78</td>
<td>57.67</td>
<td>61.56</td>
<td>67.67●</td>
<td>61.56</td>
<td>60.22●</td>
<td>57.11</td>
</tr>
<tr>
<td>Obs</td>
<td>Open</td>
<td>47.78</td>
<td>52.22</td>
<td>46.78</td>
<td>57.00</td>
<td>49.67</td>
<td>50.56</td>
<td>55.89</td>
</tr>
<tr>
<td>Self</td>
<td>Extra</td>
<td>47.78</td>
<td>48.78</td>
<td>49.67</td>
<td>57.33</td>
<td>50.56</td>
<td>54.44</td>
<td>49.89</td>
</tr>
<tr>
<td>Self</td>
<td>Emot</td>
<td>51.11</td>
<td>45.56</td>
<td>46.78</td>
<td>50.44</td>
<td>46.78</td>
<td>41.89</td>
<td>44.33</td>
</tr>
<tr>
<td>Self</td>
<td>Agree</td>
<td>52.22</td>
<td>47.89</td>
<td>50.89</td>
<td>58.33</td>
<td>56.89</td>
<td>55.22</td>
<td>52.33</td>
</tr>
<tr>
<td>Self</td>
<td>Conscl</td>
<td>51.11</td>
<td>33.44</td>
<td>45.56</td>
<td>39.33</td>
<td>43.11</td>
<td>46.11</td>
<td>53.22</td>
</tr>
<tr>
<td>Self</td>
<td>Open</td>
<td>51.11</td>
<td>52.00</td>
<td>42.22</td>
<td>61.44</td>
<td>45.00</td>
<td>56.00</td>
<td>47.78</td>
</tr>
</tbody>
</table>
Ears (speech) corpus, from observer, Naïve Bayes classifier

<table>
<thead>
<tr>
<th>Feature set</th>
<th>None</th>
<th>Type</th>
<th>LIWC</th>
<th>MRC</th>
<th>Prosody</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set size</td>
<td>0</td>
<td>4</td>
<td>88</td>
<td>14</td>
<td>11</td>
<td>All</td>
</tr>
<tr>
<td>Extraversion</td>
<td>47.78</td>
<td>45.67</td>
<td>68.89</td>
<td>68.78</td>
<td>67.56</td>
<td>73</td>
</tr>
<tr>
<td>Emotional stability</td>
<td>51.11</td>
<td>60.22</td>
<td>69.89</td>
<td>60.78</td>
<td>61.78</td>
<td>73.89</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>47.78</td>
<td>57.56</td>
<td>54.00</td>
<td>58.67</td>
<td>50.44</td>
<td>61.33</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>47.78</td>
<td>59.67</td>
<td>60.22</td>
<td>66.78</td>
<td>52.11</td>
<td>67.67</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>47.78</td>
<td>53.11</td>
<td>61.11</td>
<td>54.00</td>
<td>64.56</td>
<td>57</td>
</tr>
</tbody>
</table>
Summary

- Much easier to classifier observer-labeled than self-labeled
- Simpler classifiers like NB did well
  - not much data: 96 people, 97K words
Feature analysis: Observed Extraversion

more words
higher pitch
more concrete, imageable words
greater variation in intensity
greater mean intensity
more word repetitions
Agreeableness

- swear
  Self-assessed: Self-assessed:
- anger
  pitch variation long words, short sents
+ backchannel max intensity

<table>
<thead>
<tr>
<th>Agreeableness model with all features</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>
Conscientiousness

- swear
- anger
- negemotion

Observed:
  - +insight, +backchannel, +longwords
  - +word, +posemotion

Self-assessed:
  - +positive feelings

Conscientiousness model with all features

<table>
<thead>
<tr>
<th>#</th>
<th>Positive rules</th>
<th>α</th>
<th>#</th>
<th>Negative rules</th>
<th>α</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Occup ≥ 1.21</td>
<td>0.37</td>
<td>11</td>
<td>Swear ≥ 0.20</td>
<td>-0.18</td>
</tr>
<tr>
<td>2</td>
<td>Insight ≥ 2.15</td>
<td>0.36</td>
<td>12</td>
<td>WPS ≥ 6.25</td>
<td>-0.19</td>
</tr>
<tr>
<td>3</td>
<td>Posfeel ≥ 0.30</td>
<td>0.30</td>
<td>13</td>
<td>Pitch-mean ≥ 229</td>
<td>-0.20</td>
</tr>
<tr>
<td>4</td>
<td>Int-stddev ≥ 7.83</td>
<td>0.29</td>
<td>14</td>
<td>Othref ≥ 7.64</td>
<td>-0.20</td>
</tr>
<tr>
<td>5</td>
<td>Nlet ≥ 3.29</td>
<td>0.27</td>
<td>15</td>
<td>Humans ≥ 0.83</td>
<td>-0.21</td>
</tr>
<tr>
<td>6</td>
<td>Comm ≥ 1.20</td>
<td>0.26</td>
<td>16</td>
<td>Swear ≥ 0.93</td>
<td>-0.21</td>
</tr>
<tr>
<td>7</td>
<td>Nphon ≥ 2.66</td>
<td>0.25</td>
<td>17</td>
<td>Swear ≥ 0.17</td>
<td>-0.24</td>
</tr>
<tr>
<td>8</td>
<td>Nphon ≥ 2.67</td>
<td>0.22</td>
<td>18</td>
<td>Relig ≥ 0.32</td>
<td>-0.27</td>
</tr>
<tr>
<td>9</td>
<td>Nphon ≥ 2.76</td>
<td>0.20</td>
<td>19</td>
<td>Swear ≥ 0.65</td>
<td>-0.31</td>
</tr>
<tr>
<td>10</td>
<td>K-F-nsamp ≥ 329</td>
<td>0.19</td>
<td>20</td>
<td>Int-max ≥ 86.84</td>
<td>-0.50</td>
</tr>
</tbody>
</table>
Openness to experience

- Poor performance from Ears data – prosody helped but no language features
- But good performance from Essay data
  - Open/creative/unconventional people
    - don’t talk about school
    - use longer and rarer words
    - don’t talk about friends

<table>
<thead>
<tr>
<th>#</th>
<th>Ordered rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(School $\geq$ 1.47) and (Motion $\geq$ 1.71) $\Rightarrow$ NOT OPEN</td>
</tr>
<tr>
<td>2</td>
<td>(Occup $\geq$ 2.49) and (Sixltr $\leq$ 13.11) and (School $\geq$ 1.9) and (I $\geq$ 10.5) $\Rightarrow$ NOT OPEN</td>
</tr>
<tr>
<td>3</td>
<td>(Fam $\geq$ 600.335106) and (Friends $\geq$ 0.67) $\Rightarrow$ NOT OPEN</td>
</tr>
<tr>
<td>4</td>
<td>(Nlet $\leq$ 3.502543) and (Number $\geq$ 1.13) $\Rightarrow$ NOT OPEN</td>
</tr>
<tr>
<td>5</td>
<td>(School $\geq$ 0.98) and (You $\leq$ 0) and (AllPct $\leq$ 13.4) $\Rightarrow$ NOT OPEN</td>
</tr>
<tr>
<td>6</td>
<td>Any other feature values $\Rightarrow$ OPEN</td>
</tr>
</tbody>
</table>
Interspeech 2012 Paralinguistic challenge dataset

- SPC
- Speech clips randomly extracted from Radio Suisse Romand French news broadcasts
- 640 10-second speech clips from 322 individuals
- Emotionally neutral, no familiar words to non-French speakers
- Professional (307 samples; journalists) or nonprofessional (333 - interviewees) samples.
- Personality assessed by 11 judges
Personality labeled by BFI-10

<table>
<thead>
<tr>
<th>ID</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>This person is reserved</td>
</tr>
<tr>
<td>2</td>
<td>This person is generally trusting</td>
</tr>
<tr>
<td>3</td>
<td>This person tends to be lazy</td>
</tr>
<tr>
<td>4</td>
<td>This person is relaxed, handles stress well</td>
</tr>
<tr>
<td>5</td>
<td>This person has few artistic interests</td>
</tr>
<tr>
<td>6</td>
<td>This person is outgoing, sociable</td>
</tr>
<tr>
<td>7</td>
<td>This person tends to find fault with others</td>
</tr>
<tr>
<td>8</td>
<td>This person does a thorough job</td>
</tr>
<tr>
<td>9</td>
<td>This person gets nervous easily</td>
</tr>
<tr>
<td>10</td>
<td>This person has an active imagination</td>
</tr>
</tbody>
</table>

Extroversion: Q6 – Q1  Agreeableness: Q2 – Q7
Conscientiousness: Q8 – Q3  Neuroticism: Q9 – Q4
Openness: Q10 – Q5
### Accuracy

<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>73.5 ± 3.4</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>63.1 ± 3.7</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>71.3 ± 3.5</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>65.9 ± 3.7</td>
</tr>
<tr>
<td>Openness</td>
<td>60.1 ± 3.8</td>
</tr>
</tbody>
</table>