Lecture 11: 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/
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>
</table>
| - Yeah you would do kilograms. Yeah I see what you’re saying.  
- On Tuesday I have class. I don’t know.  
- I don’t know. A16. Yeah, that is kind of cool.  
- I don’t know. I just can’t wait to be with you and not have to do this every night, you know?  
- Yeah. You don’t know. Is there a bed in there? Well ok just... | - That’s my first yogurt experience here. Really watery. Why?  
- Damn. New game.  
- Oh.  
- That’s so rude. That.  
- Yeah, but he, they like each other. He likes her.  
- They are going to end up breaking up and he’s going to be like. |

<table>
<thead>
<tr>
<th>Unconscientious</th>
<th>Conscientious</th>
</tr>
</thead>
</table>
| - With the Chinese. Get it together.  
- I tried to yell at you through the window. 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. | - 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 [...] |
# 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 (Physcal), 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>Consc</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>Consc</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>

*Statistically significant compared to other columns.*
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></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

M5’ Regression Tree

Word count

≤ 675  > 675

Mean pitch

≤ 231  > 231

Intensity variation

≤ 6.39  > 6.39

2.86  3.02

3.23  3.83  4.24
Agreeableness

-swear  Self-assessed:
-anger    pitch variation
+backchannel  max intensity

Other-assessed:
long words, short sents

---

<table>
<thead>
<tr>
<th>Agreeableness model with all features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Positive rules</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
</tr>
<tr>
<td>1</td>
<td>Nphon $\geq$ 2.66</td>
</tr>
<tr>
<td>2</td>
<td>Tentat $\geq$ 2.83</td>
</tr>
<tr>
<td>3</td>
<td>Colon $\geq$ 0.03</td>
</tr>
<tr>
<td>4</td>
<td>Posemo $\geq$ 2.67</td>
</tr>
<tr>
<td>5</td>
<td>Voiced $\geq$ 584</td>
</tr>
<tr>
<td>6</td>
<td>Relig $\geq$ 0.43</td>
</tr>
<tr>
<td>7</td>
<td>Insight $\geq$ 2.09</td>
</tr>
<tr>
<td>8</td>
<td>Prompt $\geq$ 0.06</td>
</tr>
<tr>
<td>9</td>
<td>Comma $\geq$ 4.60</td>
</tr>
<tr>
<td>10</td>
<td>Money $\geq$ 0.38</td>
</tr>
</tbody>
</table>
Conscientiousness

- swear
- anger
- negemotion

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

Self-assessed:
- +positive feelings

<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 (Sixlitr $\leq 13.11$) 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>Trait</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>
Fun paper of the week

Durian Ripeness Striking Sound Recognition Using N-gram Models with N-best Lists and Majority Voting

Rong Phoophuangpairoj

Department of Computer Engineering, College of Engineering, Rangsit University, Thailand
rong.p@rsu.ac.th
Other datasets

- LIWC
- MRC:
- http://ota.oucs.ox.ac.uk/headers/1054.xml
Concreteness ratings

- Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.
- [http://www.humanities.mcmaster.ca/%7Evickup/Concreteness_ratings_Brysbaert_et_al_BRM.csv](http://www.humanities.mcmaster.ca/%7Evickup/Concreteness_ratings_Brysbaert_et_al_BRM.csv)
Valence, arousal, dominance


Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

http://www.humanities.mcmaster.ca/%7Evickup/Warriner_et_al_emot_ratings.csv
Age of acquisition


Supplementary data: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License.

http://www.humanities.mcmaster.ca/%7Evickup/Kuperman-BRM-data-2012.csv
Topic 2: Measuring Child-directed speech

Child-directed speech and future academic success

• By kindergarten, children from SES disadvantaged backgrounds differ in verbal and other cognitive abilities and these disparities are predictive of later academic success or failure (Hart & Risley, 1995)
  • by age 24 months, 6-month gap in language processing skills
  • between high-SES and low-SES
• Recent research suggests:
  • more talking and richer vocabulary used by parents accounts in part for these later verbal disparities.
How do we know?

- 29 Spanish-learning infants (19 and 24 months)
- At 19 months: a digital recorder in the chest pocket of specialized clothing worn by the child
- 1 day (~7 hours) of recording (selected from 1-6 days)
- LENA software produces:
  - number of adult word tokens
  - number of child vocalizations.
- Humans labeled each 5 minute segment:
  - child directed or over-heard
- Measure of child-directed speech:
  - # adult word tokens in child-directed segments/duration of the recording
LENA

Automatic Language Assessment in Three Easy Steps

1. Turn on the DLP and place it in the pocket of the child’s LENA clothing at the beginning of the recording session.

2. At the end of the recording session, plug the LENA DLP into your computer. The software automatically uploads and processes the audio file.

3. The sophisticated language environment analysis software processes the audio file and generates the LENA Reports, automatic assessments of expressive language development, and other analyses.
LENA

- segments the audio file into eight categories:
  1. adult male
  2. adult female
  3. key child
  4. other child
  5. overlapping speech
  6. noise (e.g., bumps, rattles)
  7. electronic media (e.g., radio or television)
  8. silence

- estimates # of words spoken in each adult and child segment without doing ASR
- estimates # of turns
Massive Variation in CDS: from 670 to 12,000 adult words/day!
Results

- Children who heard more child-directed speech at 19 months had **larger vocabularies** at 24 months.
- Differences in exposure to overheard speech directed to other adults and children were not related to infants’ vocabulary size.
- Amount of exposure to child-directed speech was reliably correlated with children’s **processing efficiency** at 24 months.
Processing efficiency

![Graph showing processing efficiency]

- Infants Who Heard More Child-Directed Speech at 19 Months
- Infants Who Heard Less Child-Directed Speech at 19 Months

Mean Proportion of Time Spent Looking to Target Picture at 24 Months

Time (ms) From Target-Noun Onset

0 300 600 900 1200 1500 1800