

## Supplementary Materials

---

<b>POS tags</b>	Counts of POS tags, normalized by the total number of word tokens in the sample. Includes: nouns, verbs, inflected verbs, determiners, demonstratives, adjectives, adverbs, function words, interjections, subordinate conjunctions, and coordinate conjunctions.
<b>POS tag ratios</b>	Noun: verb ratio (ratio of nouns to verbs), noun ratio (ratio of nouns to nouns and verbs), pronoun ratio (ratio of pronouns to nouns), and subordinate:coordinate ratio (ratio of subordinate conjunctions to coordinate conjunctions).
<b>Yngve depth</b>	A measure which quantifies to what extent a sentence is left-branching rather than right-branching (Yngve, 1960). We compute the maximum, mean, and total Yngve depth for each sentence, then average over all sentences in each narrative sample.
<b>Parse tree height</b>	The average height of each parse tree in the sample.
<b>Mean length of sentence (MLS)</b>	Total number of words divided by number of sentences.
<b>Mean length of clause (MLC)</b>	Total number of words divided by number of clauses (as computed by Lu’s syntactic complexity analyzer (Lu, 2010)).
<b>Mean length of T-unit (MLT)</b>	Total number of words divided by number of T-units(as computed by Lu’s syntactic complexity analyzer (Lu, 2010)). A T-unit is a minimally terminable syntactic unit consisting of a main clause and its dependent clauses.
<b>Mean word length</b>	Mean number of letters in the words in the sample.
<b>Disfluency frequencies</b>	Frequency of occurrence of the token <i>um</i> and <i>uh</i> , normalized by the total number of word tokens.
<b>“Not in dictionary” (NID) words</b>	Frequency of occurrence of word tokens of length greater than two which do not occur in the English dictionary.
<b>Total words</b>	Total number of words produced, excluding filled pauses and NID words.
<b>Type:token ratio (TTR)</b>	V/N where V is the number of word types and N is the number of word tokens.
<b>Moving-average type:token ratio (MATTR)</b>	An adaptation of TTR which reduces the effect of narrative sample length (Covington and McFall, 2010). $MATTR_w$ is the TTR calculated over a moving window of size w, and averaged over all windows.
<b>Brunets index</b>	$NV0.165$ where V is the number of word types and N is the number of word tokens (from Bucks et al. (2000) citing Brunet (1978)).
<b>Honorés statistic</b>	$100\log N/(1V_1/V)$ where $V_1$ is the number of words used only once, V is the total number of word types, and N is the number of word tokens (from Bucks et al. (2000) citing Honoré (1979)).
<b>CFG production rules</b>	The frequency of occurrence of different grammatical constituents in the data, normalized by the total number of constituents in the sample. Dependency parsing is the subject of future work.
<b>Phrase type proportion</b>	Length of each phrase type (noun phrase NP, verb phrase VP, or prepositional phrase PP), divided by total narrative length (see Chae and Nenkova (2009)).
<b>Average phrase type length</b>	Total number of words in a phrase type (noun phrase NP, verb phrase VP, or prepositional phrase PP), divided by the number of phrases of that type (see Chae and Nenkova (2009)).
<b>Phrase type rate</b>	Number of phrases of a given type (noun phrase NP, verb phrase VP, or prepositional phrase PP), divided by total narrative length (see Chae and Nenkova (2009)).
<b>Frequency</b>	Frequency with which a word occurs in some corpus of natural language, here Brysbaert and New (2009).
<b>Familiarity</b>	Subjective rating of how familiar a word seems. (Gilhooly and Logie, 1980; Stadthagen-Gonzalez and Davis, 2006).
<b>Imageability</b>	Subjective rating of how easily a word generates an image in the mind (Gilhooly and Logie, 1980; Stadthagen-Gonzalez and Davis, 2006).
<b>Age of acquisition (AOA)</b>	Subjective rating of how old a person is when they first learn that word (Gilhooly and Logie, 1980; Stadthagen-Gonzalez and Davis, 2006).
<b>Light verbs</b>	Number of occurrences of <i>be</i> , <i>have</i> , <i>come</i> , <i>go</i> , <i>give</i> , <i>take</i> , <i>make</i> , <i>do</i> , <i>get</i> , <i>move</i> , and <i>put</i> , normalized by total number of verbs (Breedin et al., 1998).
<b>Information units</b>	Binary feature that measures whether or not any of the words relating to a given information unit were mentioned (from the list of relevant information units in Croisile et al. (1996)). For example, in the sentence <i>The boy is getting a cookie and the boy is falling off the stool</i> , the feature Info unit: boy would have a value of 1.
<b>Key words</b>	Integer count of how often specific relevant words are mentioned. For example, in the sentence <i>The boy is getting a cookie and the boy is falling off the stool</i> , the key word feature for boy would have a value of 2.
<b>Cosine distance</b>	The cosine distance measures the similarity between two utterances; if they are identical, then their cosine distance is zero. The feature <i>ave_cos_dist</i> measures the average cosine distance between every pair of utterances in the transcript. The feature <i>min_cos_dist</i> measures the minimum cosine distance between pairs of utterances. We also measure the proportion of sentence pairs whose cosine distance is less than or equal to a threshold, for threshold = 0.0, 0.3, and 0.5.

---

Table 7: Text features.

---

<b>Total duration of speech</b>	Total length of all non-silent segments, in milliseconds.
<b>Phonation rate</b>	Total duration of active speech divided by the total duration of the sample (including pauses).
<b>Mean pause duration</b>	Mean length of pauses > 150 ms.
<b>Short pause count</b>	Number of pauses > 150 ms and < 400 ms.
<b>Long pause count</b>	Number of pauses $\geq$ 400 ms.
<b>Pause:word ratio</b>	Ratio of silent segments longer than 150 ms to non-silent segments.
<b>Mean/var. F0:3</b>	Mean and variance of the fundamental frequency and first three formant frequencies.
<b>Jitter</b>	Measure of the short-term variation in the pitch (frequency) of a voice.
<b>Shimmer</b>	Measure of the short-term variation in the loudness (amplitude) of a voice.
<b>Zero-crossing rate (ZCR)</b>	An approximation for average pitch of an utterance, defined as the number of sign changes along a signal, per second.
<b>Mean instantaneous power</b>	Measure related to the loudness of the voice.
<b>First autocorrelation function</b>	Mean and maximum of the first autocorrelation function.
<b>Skewness</b>	Measure of lack of symmetry in the distribution of the amplitude of a signal, associated with a tense or “creaky” voice.
<b>Kurtosis</b>	Measure of the “peakedness” of a signals amplitude, or specifically the 4th moment of its distribution.
<b>Mean recurrence period density entropy (MRPDE)</b>	Measure of periodicity of a signal. Specifically, it measures the extent to which a time series repeats in the phase space. It is similar to linear autocorrelation.
<b>Mel-frequency cepstral coefficient (MFCC) features</b>	We measure six features relating to the MFCCs: the mean, variance, skewness, and kurtosis of the energy and the first 13 MFCCs (plus their individual velocities, indicated by $\Delta$ , and accelerations, indicated by $\Delta\Delta$ ), as well as the skewness and kurtosis of their individual means.

---

Table 8: Acoustic features.

---

<b>Valence</b>	Degree of positive or negative emotion associated with a word.
<b>Arousal</b>	Intensity of the emotion associated with a word.
<b>Dominance</b>	Degree of control associated with a word.
<b>First person words</b>	Normalized occurrence count of <i>I, me, my, mine</i>
<b>Maximum voiced frequency</b>	Frequency boundary separating periodic and aperiodic components of the speech signal. (Drugman and Stylianou, 2014)
<b>Glottal closure instants</b>	Instances of significant excitation of the vocal tract.
<b>Linear prediction residuals</b>	Difference between a source-filter speech model and the observed signal (Drugman, 2014).
<b>Peak slope</b>	Parameter which differentiates breathy versus tense voice quality (Kane and Gobl, 2011).
<b>Glottal flow (and derivative)</b>	Estimate of the air flow through the vocal folds.
<b>Normalized amplitude quotient (NAQ)</b>	Parametrization of glottal closing phase (Drugman et al., 2012).
<b>Quasi-open quotient (QQQ)</b>	Parameter describing the relative open time of the glottis (Drugman et al., 2012).
<b>Harmonic richness factor</b>	Measure of the amount of harmonics in the glottal source (Drugman et al., 2012).
<b>Parabolic spectral parameter</b>	Frequency domain measure of the glottal flow (Alku et al., 1997).
<b>Cepstral peak prominence</b>	Measure of voice quality based on the cepstrum (Fraile and Godino-Llorente, 2014).

---

Table 9: New features.