Studying sentence stress using corpus data

Arto Anttila
Stanford University

Naomi Tachikawa Shapiro
University of Washington

This tutorial builds on the following article:

Anttila, Arto, Timothy Dozat, Daniel Galbraith, and Naomi Shapiro. 2020. Sentence stress in presidential speeches. In Gerrit Kentner and Joost Kremers (eds.), *Prosody in Syntactic Encoding*, Walter De Gruyter: Berlin/Boston, pp. 17-50.

Here is a link to the official website of the book which contains the *published and quotable version*:

https://www.degruyter.com/view/title/565464

Studying sentence stress using corpus data

Arto Anttila
Stanford University

Naomi Tachikawa Shapiro
University of Washington

1. Introduction

(1) Words in a sentence often differ in their relative prominence. Such prominence differences are called STRESS.

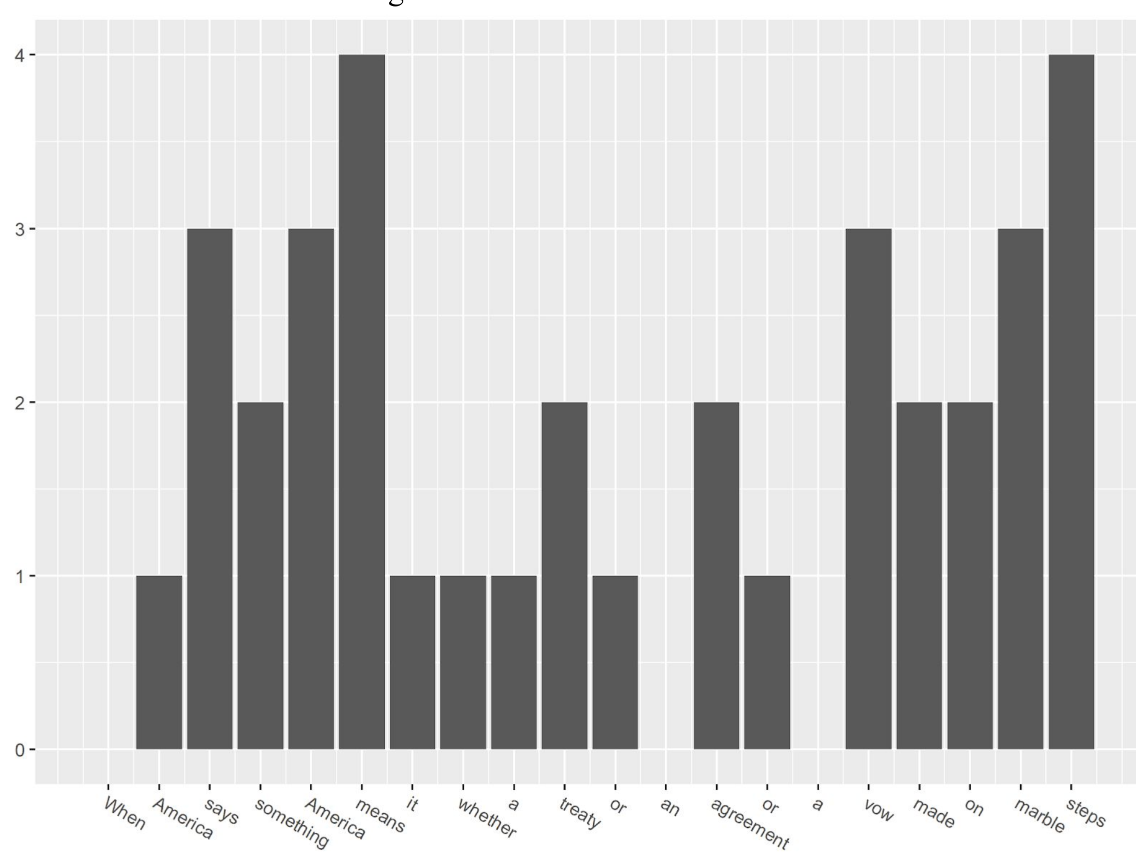
(2)

0	1	3	2		3	4	1		(Annotator 1)
0	1	3	1		2	4	1		(Annotator 2)
When	America	says	something	,	America	means	it	,	

1	1	2	1	0	2	1	0	3	2	2	3	4		(Annotator 1)
1	0	2	0	0	3	0	0	4	2	1	3	2		(Annotator 2)
whether	a	treaty	or	an	agreement	or	a	vow	made	on	marble	steps	.	

(George H. W. Bush, January 20, 1989)

(3) Both annotators heard five degrees of stress. Here's Annotator 1.



(4) Each word has a stress value from 0 to 4. Higher number means more stress.

(5) Why do we need a native speaker to hear stress?

> "[T]rained phoneticians may be able to distinguish only two levels of stress in nonsense syllable sequences but will distinguish reliably four or five levels of stress in meaningful utterances."
>
> (Halle and Keyser 1971: 17)

> "[W]hat makes a syllable accented is for the large part the work of the perceiver, generating his internal accent pattern on the basis of a strategy by which he assigns structures to the utterances. These structures, however, are not fabrications of the mind only, for they can be related to sound cues."
>
> (van Katwijk 1974: 5; cited in Baart 1987: 4)

(6) Perceived stress is a mix of at least two things (Jespersen 1920: 212-222):

a. MEANINGFUL STRESS signals emphasis, novelty, contrast, focus, etc.

He served the coffee; she needed the coffee.
(Correction: **NEEDED** the coffee.)

(*Entertainment Weekly*, November 17, 2016)

b. MECHANICAL STRESS is based on syntax and rhythm.

Q: How much did they pay you for participating in the experiment?
A: Five **FRANCS**. (Ladd 1996: 166)

(7) Phrasal stress is in many ways puzzling and mysterious. For example:

a. Why do we have stress peaks on *means* and *steps*?
b. Why do the annotators agree on *means* but disagree on *marble* and *steps*?
c. Why does the first *America* have less stress than the second *America*?

(8) Phrasal stress is in many ways puzzling and mysterious. For example:

a. Why do we have stress peaks on *means* and *steps*?
b. Why do the annotators agree on *means* but disagree on *marble* and *steps*?
c. Why does the first *America* have less stress than the second *America*?

2. Procedure

(9) Native speaker intuitions are great, but phrasal stress is a hard problem:
  a. difficult to measure phonetically
  b. varies within/across speakers
  c. native speakers sometimes disagree on the details

(10) Our approach:
  a. Procure a transcribed video and/or audio recording.
  b. Annotate the text for perceived stress. (This requires native speakers.)
  c. Annotate the text for linguistic structure using NLP tools.
  d. Implement a computational model of phrasal stress.
  e. Compare the model's predictions against the perceived stress annotations.

(11) We test a combined model of **mechanical stress** and **meaningful stress** using mixed-effects regression.

3. Perceived stress

(12) Data: Inaugural addresses of recent U.S. presidents from the American Presidency Project (Woolley and Peters 1999-, https://www.presidency.ucsb.edu/).
  a. The delivery is almost disfluency-free.
  b. The speaker is emotionally involved which expands his prosodic range.
  c. The situation is well controlled (as in Labov's department store study).

(13) Perceived stress annotations were collected through METRICGOLD (Anttila, Dozat, Galbraith, and Shapiro 2020). The source code is available on GitHub: https://github.com/tsnaomi/metric-gold.
  a. We focus on the first inaugural addresses of six U.S. presidents: Carter, Reagan, Bush Sr., Clinton, Bush Jr., and Obama.

(14) The annotators were instructed to consider the cues in the following order:
  a. clear intuitions are the best
  b. tapping, humming (annotator), head nods, hand gestures (president)
  c. explicit phonetic cues

(15) This approach to annotation purposefully draws upon <u>the entire native speaker experience</u> about stress, not just, e.g., acoustics (which we have independently available).

(16) The MetricGold web interface:

4. Mechanical stress

(17) The SPE theory (Chomsky, Halle, and Lukoff 1956; Chomsky and Halle 1968; Liberman and Prince 1977; Cinque 1993) is a good starting point:

a. The theory is simple and explicit.
b. It can be implemented and tested.

(18) The SPE phrasal stress rules:

a. Stress is assigned cyclically from the inside out, assigning `[1 stress]` to a designated word and reducing stress elsewhere by one (stress subordination).
b. **The Nuclear Stress Rule (NSR):** Assign `[1 stress]` to the rightmost vowel bearing the feature [1 stress]. Applies to phrases (NP, VP, AP, S).
c. **The Compound Stress Rule (CSR):** Skip over the rightmost word and assign `[1 stress]` to the rightmost remaining `[1 stress]` vowel; if there is no `[1 stress]` to the left of the rightmost word, then try again without skipping the word. Applies to words (N, A, V).

(19) [[[John's] [[[black] [board]] [eraser]]] [was stolen]]

```
  1        1     1      1        1        Lexical stresses
        [  1     2   ]                    Cycle 1 (CSR)
      [    1     3      2   ]             Cycle 2 (CSR)
[ 2        1     4      3    ]            Cycle 3 (NSR)
[ 3        2     5      4        1   ]    Cycle 4 (NSR)
```

(20) <u>Computational model</u>: MetricalTree (Anttila, Dozat, Galbraith, and Shapiro 2020) implements these rules in Python. MetricalTree uses syntactic parses provided by the Stanford Parser (Klein and Manning 2003; Chen and Manning 2014; Manning et al. 2014). The source code for MetricalTree is available at https://github.com/tdozat/Metrics

(21) Function words are a major puzzle in English prosody (see, e.g., Selkirk 1996). Are *a*, *all*, *in*, *is*, *not*, *that*, *the*, *this*, *will*, *you*, etc., stressed or unstressed? This decision will impact the calculation of phrasal stress in virtually every sentence (Altenberg 1987; Hirschberg 1993; Nenkova et al. 2007; Shih 2014, Anttila and Heuser 2016; Anttila 2017; Anttila; Dozat, Galbraith and Shapiro 2020, a.o.).

(22) <u>Working hypothesis</u>: Words are always lexically stressed except for those listed below:

Unstressed	Stress-ambiguous
it, *to*, determiners, complementizers, coordinating conjunctions, possessive pronouns, expletive nominals, interjections	*this*, *that*, *these*, *those*, prepositions, modals, personal pronouns, pre-determiners, copulas, auxiliaries, negation words, *wh*-words

(23) The following sentence has 15 stress-ambiguous words and thus 32,762 possible stress paths (Obama 2009). MetricalTree annotates the words as **ambiguous**, stressed, or unstressed.

> To **those who** cling to power **through** corruption and deceit and the silencing **of** dissent , know **that you are on** the wrong side **of** history , but **that we will** extend a hand **if you are** willing to unclench your fist .

(24) Lexical stress ambiguities must be resolved before phrasal stress rules can apply. We consider four different stress mechanical stress models:

a. Model 1: Ambiguous words are stressed.
b. Model 2: Ambiguous monosyllables are unstressed, polysyllables stressed.
c. Model 3: Ambiguous words are unstressed.
d. Ensemble model that takes the mean of the three prior models.

(25) Mechanical stress (MetricalTree, Model 2) converted to grids:

```
                     x
                x
x     x         x
x     x   x     x
Sandy voted for Kim.
SPE: 2 2 3 1
```

```
      x
x     x
x     x
x     x   x
x     x   x   x
Sandy voted for him.
SPE: 2 1 4 5
```

```
              x
    x         x
    x     x   x
x   x     x   x
She voted for Kim.
SPE: 4 2 3 1
```

5. Meaningful stress

(26) "What is informative is accented, what is uninformative is unaccented" (Bolinger 1972: 634; Pan and McKeown 1999; Pan and Hirschberg 2000; Cole et al. 2010).

(27) We quantify a word's *informativity* to assess meaning-driven stress ("meaningful stress").

(28) We calculate the INFORMATIVITY of a word w as the weighted average of the negative log predictability of seeing w given each context c that w follows in the corpus, beginning with Roosevelt 1933 (cf. Piantadosi et al. 2011; Cohen Priva 2012, 2015):

$$\text{inform}(w) = -\sum_{c \in C} P(c|w) \log_2 P(w|c)$$

a. If c is the directly preceding word → *bigram informativity*
b. If c is the two preceding words → *trigram informativity*

(29) *Informativity* (*of*) = 1.7024
Informativity (*Americans*) = 5.4126
Informativity (*confronting*) = 6.0549

6. Creating a data frame

(30) We compile all of the relevant information in a data frame. Each row in the data frame represents one token (word or punctuation), coupled with the perceived stress annotation from one annotator (i.e., the total number of rows = the number of tokens × the number of annotators):

word	seg	nseg	lexstress	pos	category	sidx	norm_m1
My	M AY1	2	no	PRP$	FUNC	1	0.4
fellow	F EH1 L OW0	4	yes	JJ	ADJ	1	0.6
citizens	S IH1 T AH0 Z AH0 N Z	8	yes	NNS	NOUN	1	0.8
,		NA	???	,	NA	1	NA
I	AY1	1	ambig	PRP	FUNC	1	0.6
stand	S T AE1 N D	5	yes	VBP	VERB	1	0.6
here	HH IH1 R	3	yes	RB	NA	1	0.4
today	T AH0 D EY1	4	yes	NN	NOUN	1	0.6
humbled	HH AH1 M B AH0 L D	7	yes	VBN	VERB	1	0.6
by	B AY1	2	ambig	IN	FUNC	1	0.4
the	DH AH0	2	no	DT	FUNC	1	0.2
task	T AE1 S K	4	yes	NN	NOUN	1	0.6
before	B IH0 F AO1 R	5	ambig	IN	FUNC	1	0.4
us	AH1 S	2	ambig	PRP	FUNC	1	0.6

(31) Phonological, syntactic, and mechanical stress annotations by MetricalTree:

president	the president who gave the inaugural speech
widx	word index
norm_widx	word index normalized (= widx / sentence length)
word	English spelling
seg	ARPABET transcription
nseg	number of segments in the word
lexstress	lexical stress (yes, no, ambig)
pos	part of speech
category	part of speech category (NOUN, ADJ, VERB, FUNC)
sidx	sentence index
sent	the whole sentence
m1	mechanical stress Model 1
m2a	mechanical stress Model 2
m2b	mechanical stress Model 3
mean	mechanical stress Ensemble Model
norm_m1	mechanical stress Model 1 normalized
norm_m2a	mechanical stress Model 2 normalized
norm_m2b	mechanical stress Model 3 normalized
norm_mean	mechanical stress Ensemble Model normalized

(32) Statistics derived from the inaugural corpus (in particular, frequency and informativity):

doc.freq	the frequency of the word in the document
d.inform.2	document-level bigram informativity
d.inform.3	document-level trigram informativity
corpus.freq	the frequency of the word across the corpus
c.inform.2	corpus-level bigram informativity
c.inform.3	corpus-level trigram informativity

(33) Perceived stress annotations collected through MetricGold:

annotator	annotator identifier
perc	perceived stress
norm_perc	perceived stress normalized
UB	utterance boundary
UF	utterance-final word

(34) <u>Stress normalization</u>: Normalized *predicted* values are those named `norm_[model]`, where each predicted stress is divided by the maximum stress predicted for that sentence by the given model. Likewise, normalized *perceived* values are listed under `norm_perc`.

```
                                        x
                              x    x    x
      x                       x    x    x          (Ronald Reagan 1981,
      x      x     x          x    x    x           raw grid, Model 2)
  x   x      x     x          x    x    x
The business of    our   nation  goes forward
```

Column	Sentence							Content
sent	*The*	*business*	*of*	*our*	*nation*	*goes*	*forward*	*the text*
norm_m1	0.2	0.6	0.6	0.4	0.8	0.8	1	SPE, norm
norm_m2a	0.2	0.6	0.4	0.4	0.8	0.8	1	SPE, norm
norm_m2b	0.2	0.6	0.4	0.4	0.8	0.8	1	SPE, norm
norm_mean	0.2	0.6	0.467	0.4	0.8	0.8	1	SPE, norm
norm_perc$_{A1}$	0	1	0	0	0.333	0.333	0.667	perc, norm
norm_perc$_{A2}$	0	0.667	0	0	0.333	0	1	perc, norm

7. Visualizing the data in R

(35) Install R (R Core Team 2019) and the following packages:

```
library(languageR)    # Baayen and Shafaei-Bajestan 2019
library(lme4)         # Bates et al. 2015
library(lmerTest)     # Kuznetsova et al. 2017
library(MASS)         # Venables and Ripley 2003
library(ggplot2)      # Wickham 2016
library(mgcv)         # Wood 2017
```

(36) Tip: Check out our GitHub repository! The data frame `inaugurals.csv` contains the inaugural addresses for Bush Jr. (2001) and Obama (2009), plus the R script for this tutorial:

https://github.com/tsnaomi/AMP-2020-Tutorial

(37) Read the spreadsheet into R and name it something you can remember. Here, we choose `pres6`, since we focus on six inaugural addresses:

```
pres6 <- read.csv("inaugurals.csv", header = T)
```

or

```
pres6 <- read.csv(file.choose(), header = T)
```

(38) We focus on adjectives, nouns, verbs, and function words, trimming the data down from 11,641 to 10,982 words.

(39) We start by visualizing the following relationships:

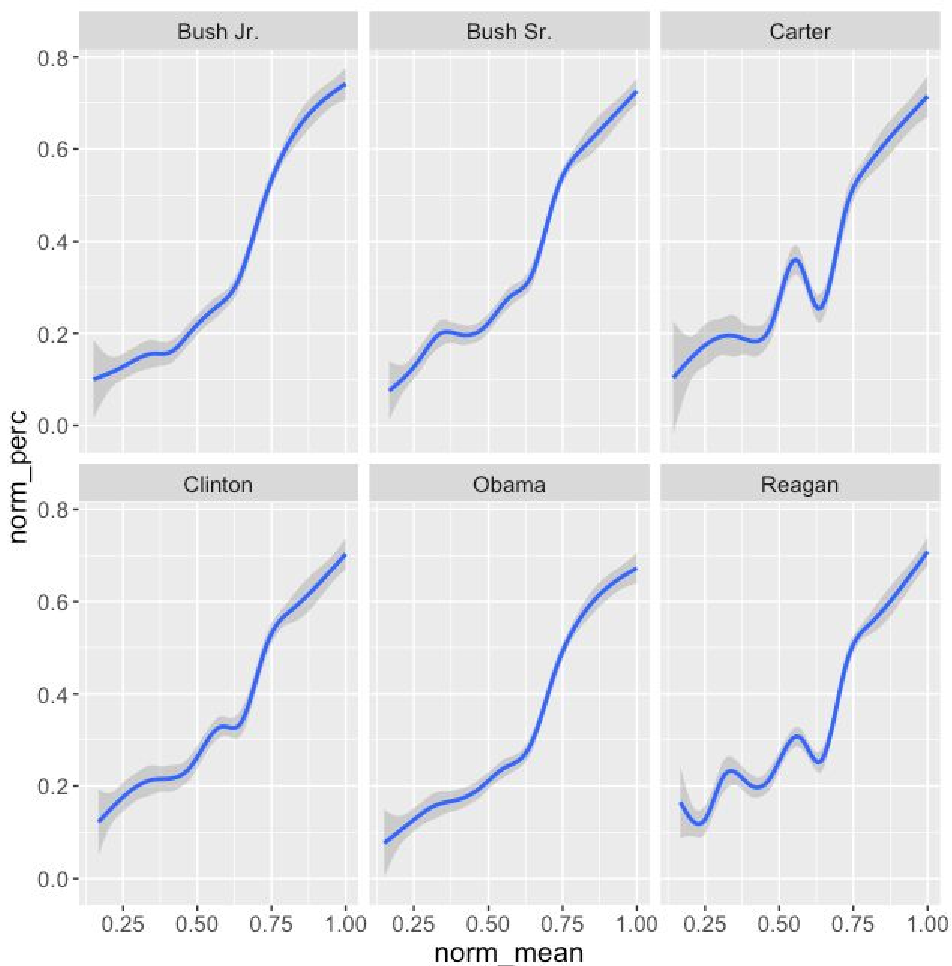
a. Mechanical stress vs. perceived stress
   How well do the NSR and CSR predict perceived stress?
b. Informativity/frequency vs. perceived stress
   Are more informative and less frequent words more stressed?
c. Mechanical stress vs. informativity
   Are informative words placed in mechanically stressed positions?
d. Parts of speech
   How do nouns, adjectives, verbs, and function words differ?

7.1 Mechanical stress vs. perceived stress

(40) How well do the SPE stress rules work?

x-axis: Mechanical stress (Ensemble Model, normalized) from no stress (0.00) to high stress (= 1.00)
y-axis: Perceived stress (normalized) from soft (= 0.0) to loud (= 1.0)

```
ggplot(pres6, aes(x=norm_mean, y=norm_perc)) +
     geom_smooth() +
     facet_wrap(~president)
```



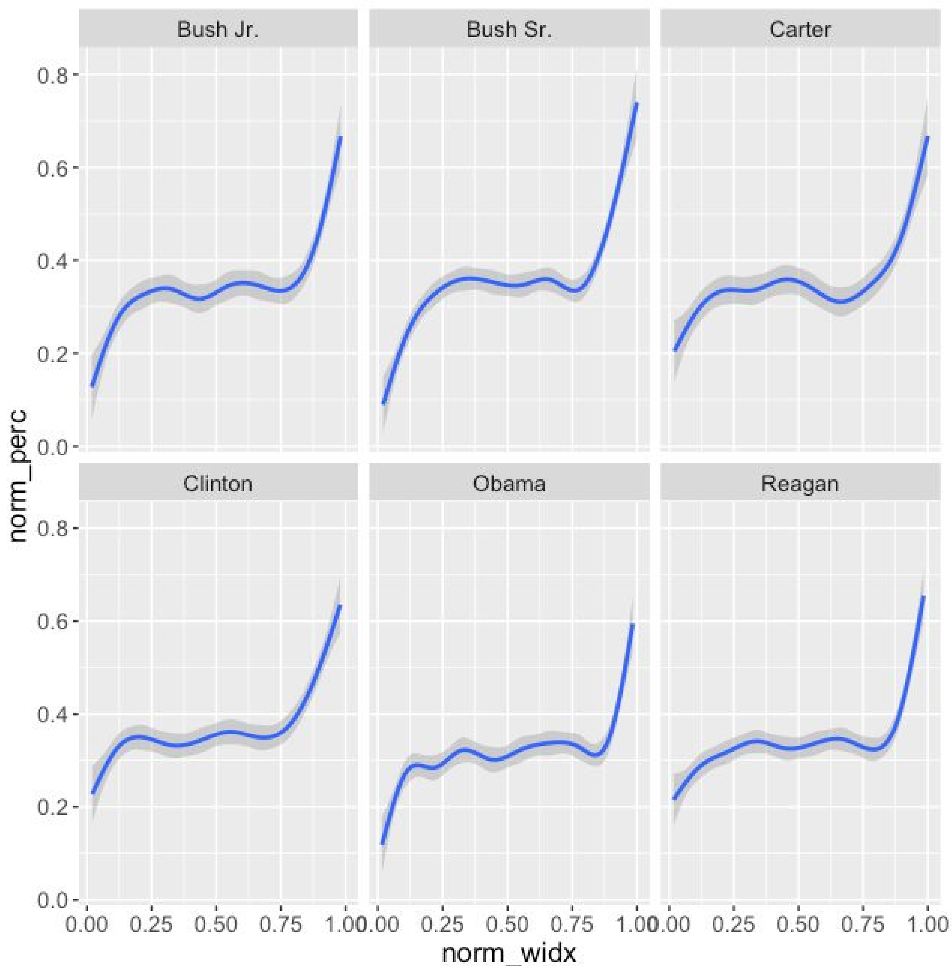
(41) Overlaid on top of the scatterplot (data points not shown) is a smooth curve that includes an assessment of uncertainty in the form of pointwise confidence intervals, shown in grey. By default, ggplot2 fits a generalized additive model provided by the `mgcv` package (Wood 2017) when there are more than 1000 points (Wickham 2016: 19).

(42) Discovery: The SPE rules (NSR & CSR) are clearly on the right track.

(43) Does perceived stress depend on the word's linear position in the sentence?

x-axis: word position (normalized) from the beginning of the sentence (= 0.00) to the end (= 1.00)
y-axis: perceived stress (normalized) from soft (= 0.0) to loud (1.0)

```
ggplot(pres6, aes(x=norm_widx, y=norm_perc)) +
      geom_smooth() +
      facet_wrap(~president)
```



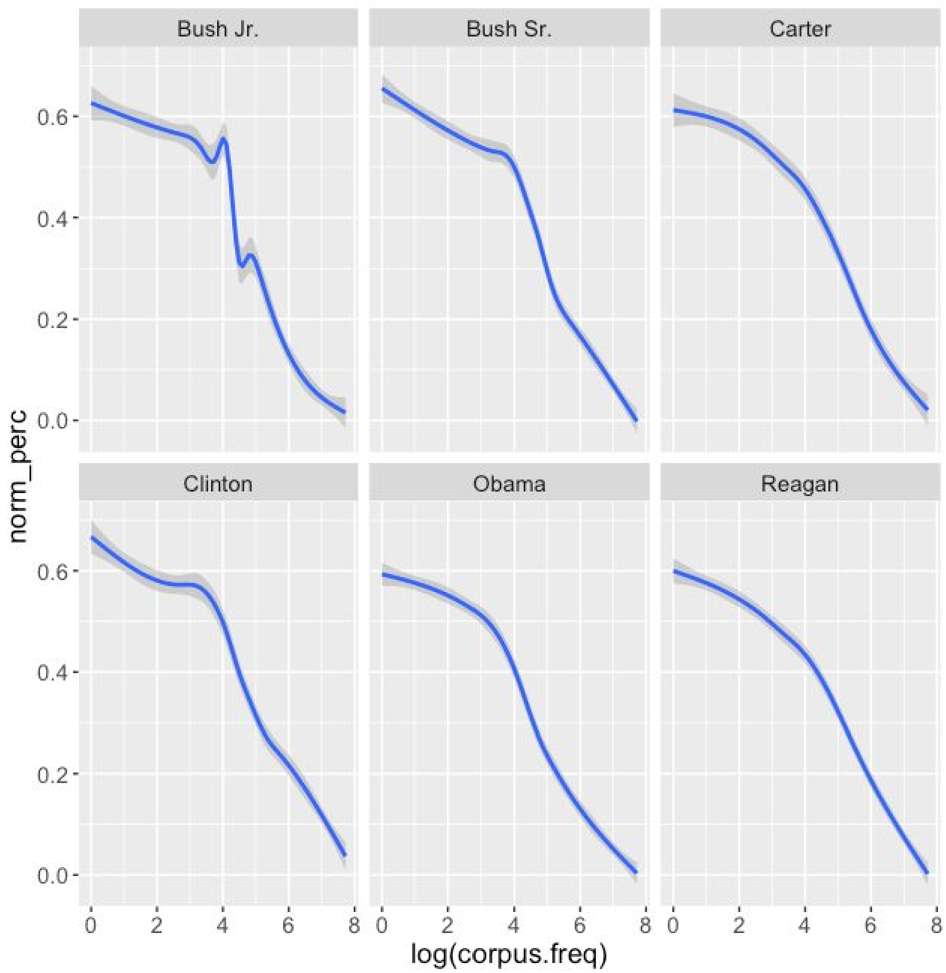
(44) Discovery: The level of perceived stress increases from the beginning of the sentence to the end. This is consistent with the view that English phrasal stress is right-headed and cyclic.

7.2 Frequency/informativity vs. perceived stress

(45) Does a word's lexical frequency matter to the perception of stress?

<u>x-axis</u>: word frequency (logged) in the presidential corpus, beginning from Roosevelt 1933
<u>y-axis</u>: perceived stress (normalized) from soft (= 0.0) to loud (1.0)

```
ggplot(pres6, aes(x=log(corpus.freq), y=norm_perc)) +
     geom_smooth() +
     facet_wrap(~president)
```

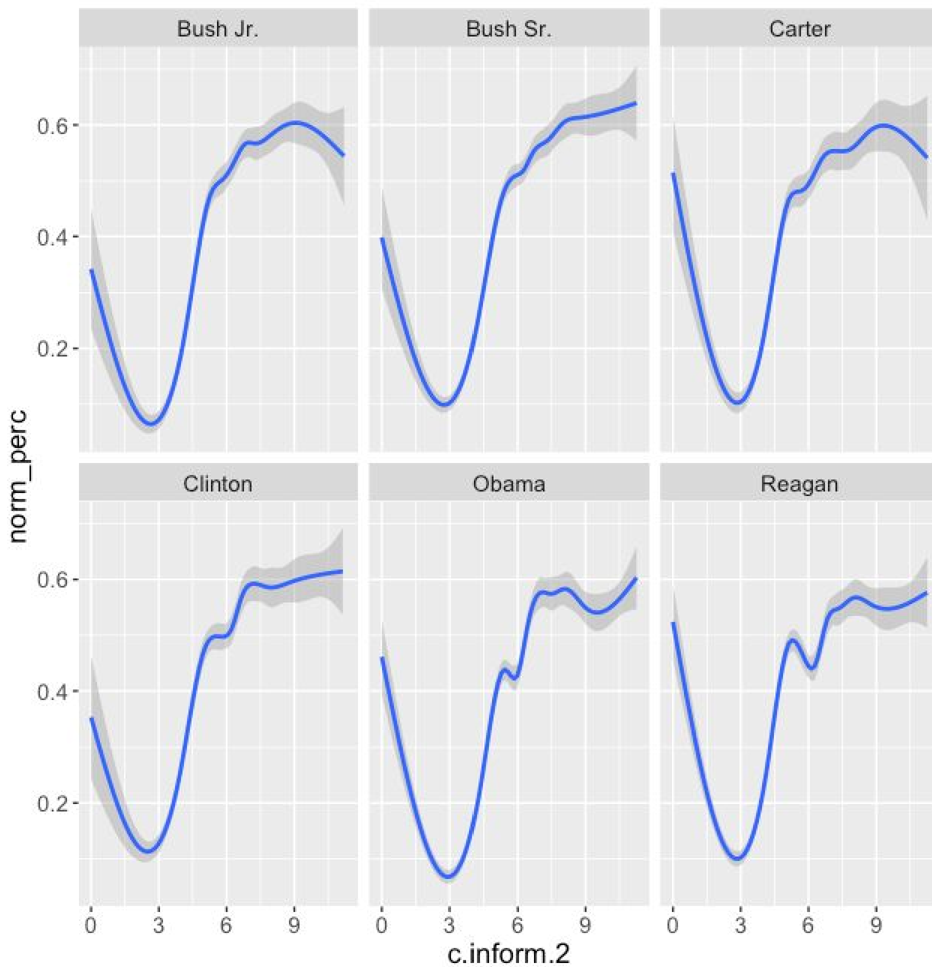


(46) <u>Discovery</u>: The more frequent the word, the less it is stressed.

(47) Bolinger (1972) made the following claim: "What is informative is accented, what is uninformative is unaccented" (Bolinger 1972: 634). Is that so?

x-axis: Bigram informativity of a word
y-axis: Perceived stress (normalized) from soft (= 0.0) to loud (1.0)

```
ggplot(pres6,aes(x=c.inform.2, y=norm_perc)) +
      geom_smooth() +
      facet_wrap(~president)
```



(48) Discovery: The word's bigram informativity and its perceived stress are indeed positively correlated, but only in words with informativity > 3.

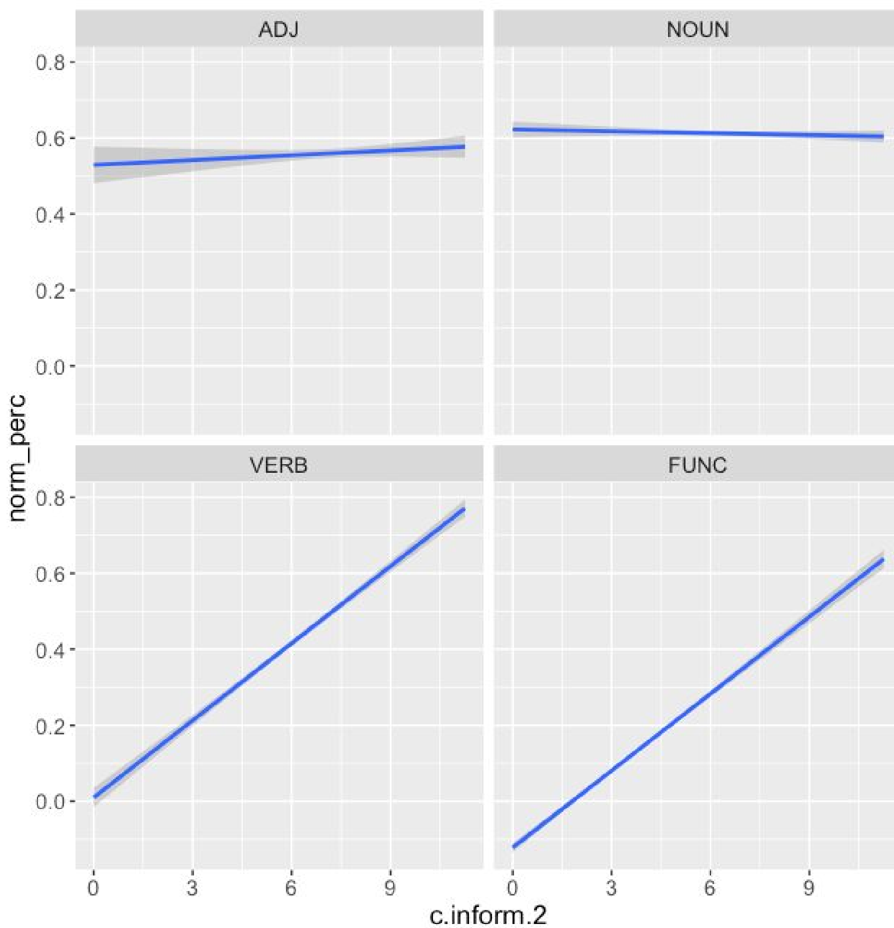
(49) Puzzle: There's something strange going on when informativity < 3. What is it? Is there an error in the dataframe?

(50) How about if we split the data by part of speech (instead of by president)?

(51) Here's the same informativity vs. perceived stress correlation by part of speech that you can get by `facet_wrap(~category)`. This looks more interesting!

<u>x-axis</u>: Bigram informativity of a word
<u>y-axis</u>: Perceived stress (normalized) from soft (= 0.0) to loud (1.0)

```
ggplot(pres6, aes(x=c.inform.2, y=norm_perc)) +
    geom_smooth(method = "lm") +
    facet_wrap(~category)
```



(52) This time, we used `method = "lm"`, which fits a linear model, giving the line of best fit (Wickham 2019: 20).

(53) <u>Discovery</u>: The "Bolinger effect" is found only amongst verbs and function words, not amongst nouns and adjectives.
  a. Perceived stress tells us nothing about the informativity of a noun.
  b. This suggests stress on nouns is entirely mechanical.

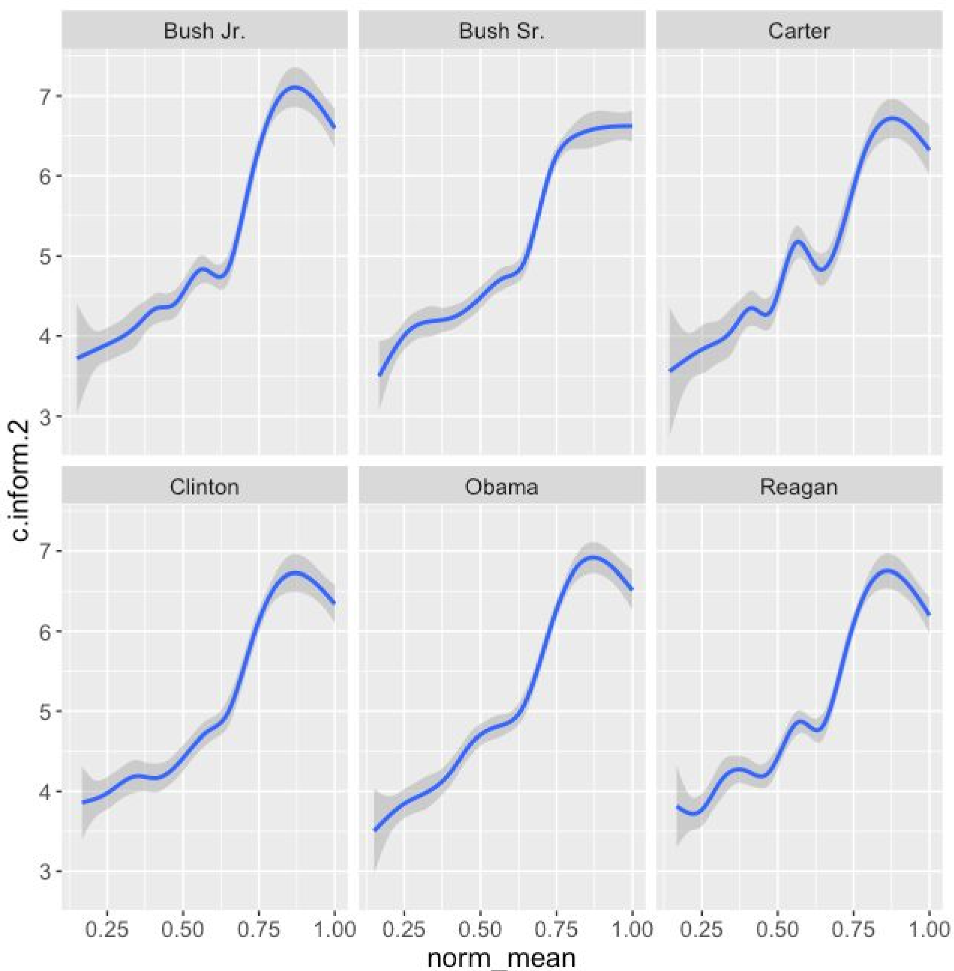
7.3 Mechanical stress vs. informativity

(54) Bolinger (1957) hypothesized that informative material tends to be placed where mechanical stress is highest, at least in good prose. We might call this hypothesis STRESS-INFORMATION ALIGNMENT.

(55) Mechanical stress vs. bigram informativity:

x-axis: Mechanical stress (Ensemble Model, normalized) from no stress (0.00) to high stress (= 1.00)
y-axis: Bigram informativity of a word

```
ggplot(pres6, aes(x=norm_mean, y=c.inform.2)) +
     geom_smooth() +
     facet_wrap(~president)
```



(56) Discovery: The data are consistent with Bolinger's (1957) hypothesis.

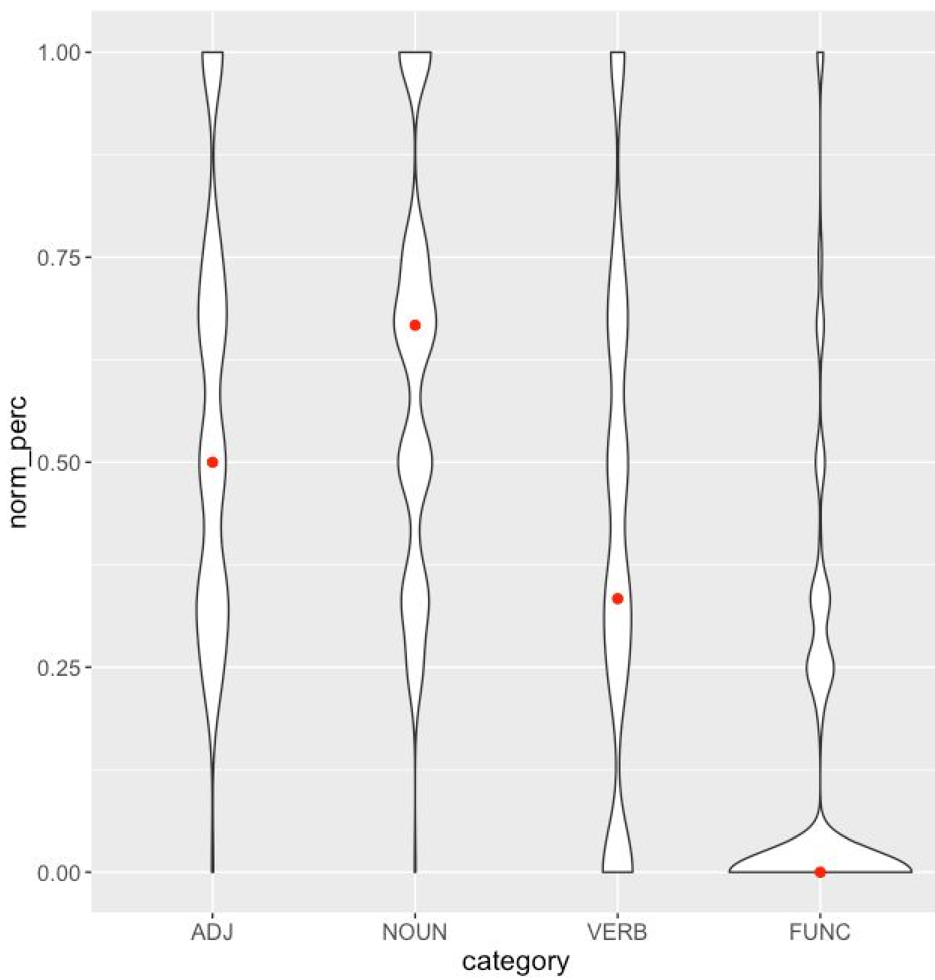
7.4 Part-of-speech effects

(57) It is known that nouns are more faithful than adjectives, which are more faithful than verbs, in word phonology (Smith 2011). Similar "accentability hierarchies" have been found in sentence phonology (Ladd 1980: 84-6' see Baart 1987: 56-57 for a summary; Pan and Hirschberg 2000: 237; Pan et al. 2002). Can we find evidence for the differential accountability of parts of speech in our data?

x-axis: Parts of speech (function word, verb, adjective, noun)
y-axis: Perceived stress (normalized) from soft (= 0.0) to loud (1.0)

(The red dot is the median.)

```
ggplot(pres6, aes(category, norm_perc)) +
      geom_violin() +
      stat_summary(fun=median, geom="point",
                   size=2, color="red")
```



(58) Discovery: Function words are the softest, nouns are loudest.

8. Regression modeling

(59) A linear regression model:

Response:	norm_perc	Perceived stress (normalized)
Predictors:	log(corpus.freq)	The word's corpus frequency (logged)
	norm_mean	SPE stress (Ensemble Model, normalized)
	category	The word's syntactic category
	norm_widx	The word's sentence position (normalized)
	log(nseg)	The word's length in segments (logged)

(60) All five predictors are independently significant:

```
presidents.fit = lm(
     perc ~ norm_mean + log(corpus.freq) + log(nseg) +
          norm_widx + category,
     data = pres6
     )

summary(presidents.fit)
```

Output:

```
Call:
lm(formula = perc ~ norm_mean + log(corpus.freq) + log(nseg) +
    norm_widx + category, data = pres6)

Residuals:
    Min      1Q  Median      3Q     Max
-2.8008 -0.5435 -0.1968  0.4386  4.6933

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       1.728432   0.043665  39.584  < 2e-16 ***
norm_mean         0.413989   0.043032   9.620  < 2e-16 ***
log(corpus.freq) -0.162300   0.004153 -39.083  < 2e-16 ***
log(nseg)         0.183966   0.017172  10.713  < 2e-16 ***
norm_widx         0.191281   0.023790   8.040 9.40e-16 ***
categoryFUNC     -0.701559   0.027071 -25.915  < 2e-16 ***
categoryNOUN      0.117097   0.024083   4.862 1.17e-06 ***
categoryVERB     -0.374526   0.025417 -14.735  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8498 on 21822 degrees of freedom
  (4546 observations deleted due to missingness)
Multiple R-squared:  0.4808,   Adjusted R-squared:  0.4806
F-statistic:  2887 on 7 and 21822 DF,  p-value: < 2.2e-16
```

(61) A mixed effects linear regression model (Bates et al. 2014; Kuznetsova et al. 2016) with president and annotator as random intercepts:

```
presidents.lmer = lmer(
     perc ~ norm_mean + log(corpus.freq) + log(nseg) +
          norm_widx + category +
          (1|president) + (1|annotator),
     data = pres6
     )

summary(presidents.lmer)
```

Output:

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method
['lmerModLmerTest']
Formula: perc ~ norm_mean + log(corpus.freq) + log(nseg) + norm_widx +
    category + (1 | president) + (1 | annotator)
   Data: pres6

REML criterion at convergence: 53885.9

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.4968 -0.6391 -0.1862  0.4787  5.6441

Random effects:
 Groups    Name        Variance Std.Dev.
 president (Intercept) 0.01410  0.1188
 annotator (Intercept) 0.04476  0.2116
 Residual              0.68878  0.8299
Number of obs: 21830, groups:  president, 6; annotator, 2

Fixed effects:
                   Estimate Std. Error         df t value Pr(>|t|)
(Intercept)       1.759e+00  1.630e-01  1.401e+00  10.796   0.0256 *
norm_mean         4.115e-01  4.204e-02  2.182e+04   9.788  < 2e-16 ***
log(corpus.freq) -1.662e-01  4.063e-03  2.182e+04 -40.914  < 2e-16 ***
log(nseg)         1.773e-01  1.682e-02  2.182e+04  10.538  < 2e-16 ***
norm_widx         1.923e-01  2.323e-02  2.182e+04   8.279  < 2e-16 ***
categoryFUNC     -6.804e-01  2.648e-02  2.182e+04 -25.695  < 2e-16 ***
categoryNOUN      1.305e-01  2.356e-02  2.182e+04   5.538 3.09e-08 ***
categoryVERB     -3.623e-01  2.485e-02  2.182e+04 -14.579  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) nrm_mn lg(c.) lg(ns) nrm_wd ctFUNC ctNOUN
norm_mean   -0.108
lg(crps.fr) -0.140  0.053
log(nseg)   -0.165 -0.151  0.449
norm_widx   -0.011 -0.468 -0.003  0.053
categryFUNC -0.123  0.178 -0.355  0.151 -0.043
categryNOUN -0.071 -0.180 -0.060 -0.057  0.091  0.614
categryVERB -0.107 -0.167 -0.031  0.163  0.127  0.669  0.717
```

9. Summary

(62) a. Perceived phrasal stress is influenced by multiple factors.
b. Syntax and the NSR & CSR play an important role in phrasal stress.
c. Nominal stresses (nouns, adjectives) are loud and meaningless.

Acknowledgements

Financial support: Office of the Vice-Provost for Undergraduate Education, Stanford University, and the Roberta Bowman Denning Initiative Committee, H&S Dean's Office. Other support: Jared Bernstein, Canaan Breiss, Joan Bresnan, Uriel Cohen Priva, Joshua Falk, Vivienne Fong, Carlos Gussenhoven, Penny Eckert, Isabelle Franz, Bruce Hayes, Ryan Heuser, Larry Hyman, Gerrit Kentner, Paul Kiparsky, Mark Liberman, Geoffrey Pullum, Nathan Sanders, Madeline Snigaroff, Benjamin Storme, Meghan Sumner, Simon Todd, Alex Wade.

Data source

Woolley, John T. and Gerhard Peters. 1999--. The American Presidency Project, UC Santa Barbara, https://www.presidency.ucsb.edu/

References

Altenberg, Bengt. 1987. *Prosodic Patterns in Spoken English: Studies in the Correlation Between Prosody and Grammar for Text-to-Speech Conversion*. Vol. 76. Studentlitteratur, Lund.

Anttila, Arto. 2017. Stress, phrasing, and auxiliary contraction in English. In Vera Gribanova and Stephanie Shih (eds.), *The Morphosyntax-Phonology Connection: Locality and Directionality at the Interface*, Oxford University Press. pp. 143-170.

Anttila, Arto & Ryan Heuser. 2016. Phonological and metrical variation across genres, *Proceedings of 2015 Annual Meeting on Phonology*, Linguistic Society of America, Washington DC. https://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/3679

Anttila, Arto, Timothy Dozat, Daniel Galbraith & Naomi Shapiro. 2020. Sentence stress in presidential speeches. In Gerrit Kentner and Joost Kremers (eds.), *Prosody in Syntactic Encoding*, Walter De Gruyter: Berlin/Boston, pp. 17-50.

Baart, Joan. 1987. *Focus, Syntax, and Accent Placement*. Ph.D. dissertation, University of Leiden.

Baayen, R. H. & Elnaz Shafaei-Bajestan. 2019. *languageR: Analyzing Linguistic Data: A Practical Introduction to Statistics*. R package version 1.5.0. https://CRAN.R-project.org/package=languageR.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Bolinger, Dwight L. 1957. Maneuvering for Stress and Intonation. *College Composition and Communication*, 8(4), 234-238.

Bolinger, Dwight L. 1972. Accent is predictable (if you are a mind reader). *Language* 48, 633-644.

Calhoun, Sasha. 2010. How does informativeness affect prosodic prominence? *Language and Cognitive Processes* 25(7-9), 1099-1140.

Chen, Danqi & Christopher Manning. 2014. A Fast and Accurate Dependency Parser Using Neural Networks. In *Proceedings of EMNLP* 2014.

Chomsky, Noam, Morris Halle & Fred Lukoff. 1956. On accent and juncture in English. In Morris Halle et al. (eds.), *For Roman Jakobson: Essays on the occasion of his sixtieth birthday*, Mouton & Co., The Hague, pp. 65-80.

Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Harper and Row, New York.

Cinque, Guglielmo. 1993. A null-theory of phrase and compound stress. *Linguistic Inquiry* 24, 239-298.

Cohen Priva, Uriel. 2012. *Sign and Signal: Deriving Linguistic Generalizations from Information Utility*. Ph.D. dissertation, Stanford University.

Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology* 6(2), 243-278.

Cole, Jennifer, Timothy Mahrt & Joseph Roy. 2017. Crowd-sourcing prosodic annotation. *Computer Speech & Language* 45, 300-325.

Cole, Jennifer, Yoonsook Mo & Mark Hasegawa-Johnson. 2010. Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1(2), 425-452.

Gussenhoven, Carlos. 2011. Sentential prominence in English. *The Blackwell Companion to Phonology*, Vol. 2, 2778-2806.

Hasegawa-Johnson, Mark, Ken Chen, Jennifer Cole, Sarah Borys, Sung-Suk Kim, Aaron Cohen, Tong Zhang, Jeung-Yoon Choi, Heejin Kim, Taejin Yoon & Sandra Chavarria. 2005. Simultaneous recognition of words and prosody in the Boston University radio speech corpus. *Speech Communication* 46(3-4), 418-439.

Hirschberg, Julia. 1993. Pitch accent in context: predicting intonational prominence from text. *Artificial Intelligence* 63, 305-340.

Halle, Morris & Samuel Jay Keyser. 1971. *English stress: Its form, Its Growth, and Its Role in Verse*. New York, Evanston & London: Harper & Row, Publishers.

Jespersen, Otto. 1920. *Lehrbuch der Phonetik: Mit 2 Tafeln*. BG Teubner.

Klein, Dan & Christopher D. Manning. 2003. Accurate Unlexicalized Parsing. *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, pp. 423-430.

Kuznetsova A, Brockhoff PB & Christensen RHB (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13), 1-26. doi: 10.18637/jss.v082.i13 (URL: https://doi.org/10.18637/jss.v082.i13).

Ladd, D. Robert. 1980. *The Structure of Intonational Meaning*. Indiana University Press: Bloomington & London.

Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge University Press.

Liberman, Mark & Alan Prince. On stress and linguistic rhythm. *Linguistic Inquiry* 8(2), 249-336.

Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard & David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.

Nenkova, Ani, Jason Brenier, Anubha Kothari, Sasha Calhoun, Laura Whitton, David Beaver & Dan Jurafsky. 2007. To memorize or to predict: Prominence labeling in conversational speech. *Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, http://www.aclweb.org/anthology/N07-1002. Also posted at ScholarlyCommons http://repository.upenn.edu/cis_papers/732.

Pan, Shimei & Julia Hirschberg. 2000. Modeling local context for pitch accent prediction. In *Proceedings of the 38th Annual Meeting on Association for Computational linguistics*, Association for Computational Linguistics, pp. 233-240.

Pan, Shimei & Kathleen R. McKeown. 1999. Word informativeness and automatic pitch accent modeling. In 1999 *Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, available at http://www.aclweb.org/anthology/W99-0619.

Pan, Shimei, Kathleen McKeown & Julia Hirschberg. 2002. Exploring features from natural language generation for prosody modeling. *Computer Speech & Language* 16(3-4), 457-490.

Piantadosi, Steven T., Harry Tily & Edward Gibson. 2011. Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences* 108(9), 3526-3529.

Prince, Alan. 1983. Relating to the grid. *Linguistic Inquiry* 14(1), 19-100.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Schmerling, Susan F. 1976. *Aspects of English Sentence Stress*. Texas University Press, Austin.

Shih, Stephanie. 2014. Towards Optimal Rhythm. Ph.D. dissertation, Stanford University.

Smith, Jennifer L. 2011. Category-specific effects. *The Blackwell Companion to Phonology* 4, 2439-2463.

Venables, W. N. & Ripley, B. D. 2002. *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Wagner, Michael. 2005. Asymmetries in Prosodic Domain Formation, in Norvin Richards and Martha McGinnis (eds.), *Perspectives on Phases*, MITWPL 49, 329-367.

Weide, Robert. 1998. The Carnegie Mellon Pronouncing Dictionary [cmudict.0.6], ftp://ftp.cs.cmu.edu/project/speech/dict/, Carnegie Mellon University.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wood, S.N. 2017. *Generalized Additive Models: An Introduction with R* (2nd edition). Chapman and Hall/CRC.