Frequency effects and prosodic boundary strength

Tina Bögel and Alice Turk

University of Konstanz // University of Edinburgh

ICPhS 2019

Melbourne

Overview

- The *Smooth Signal Redundancy Hypothesis*
- Experiment on frequency effects with respect to
  - the placement of syntactic boundaries
  - the strength of the resulting prosodic boundaries

  → durational measurements of the boundary-related intervals

The smooth signal redundancy hypothesis (SSRH)

- Often no reliable cues to indicate prosodic boundaries in spoken language
- **Hypothesis**: prosodic boundary structure is planned to achieve SSR

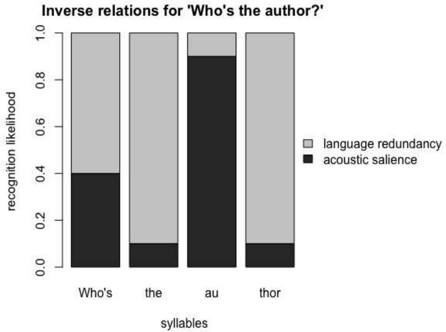
→ make the recognition of each word in an utterance equally likely

→ prosodic boundary strength assumed to **inversely** relate to language redundancy, i.e., non-acoustic information:

  - likelihood of syntactic structure
  - lexical word frequency
  - word bigram frequency
  - ...

- More predictable elements require "less explicit signal information" than less predictable elements for successful recognition (Lindblom 1990)

Inverse relation



(Aylett 2000, Aylett and Turk 2004)

- → Inverse, complementary relationship between language redundancy and acoustic redundancy
- → Recognition likelihood spread evenly throughout an utterance
- ⇒ achieve maximal understanding with minimal effort

Previous work

Previous work showed that increased

- lexical frequency (e.g., Jurafsky et al. 2001)
- bigram frequency (e.g., Aylett 2000, Aylett and Turk 2004, 2006, Pluymaekers et al. 2005, Bell et al. 2009)
- syntactic predictability (e.g., Gahl and Garnsey 2004, Watson et al. 2006)

led to a reduction of word/segment duration, and influenced the placement of syntactic boundaries.

Clearly demarcating word boundaries → more salience

Hypothesis

- Inverse relationship between language redundancy and acoustic salience should hold for prosodic boundaries
- SSRH predicts greater final lengthening, initial lengthening, initial strengthening, F0 reset, etc., given low language redundancy

→ Stronger prosodic boundaries are expected to occur where language redundancy is low, e.g., within infrequent stretches of speech

- SSRH would further predict a (gradient) correlation between boundary strength and language redundancy (e.g. greater final lengthening, initial lengthening, initial strengthening, F0 reset, etc.),

→ Has not been tested experimentally!

Work presented here

- Investigates the relationship between language redundancy and prosodic boundary strength
- through the effect of:
  - syntactic frequency
  - word frequency
  - word bigram frequency

→ on the placement of intonational phrase boundaries

→ on durational measurements of boundary strength

Challenge:

Need to vary language redundancy, while using controlled material

- with similar syntactic phrasing
- with similar segments across boundaries (effects might be subtle)

Experimental design: syntactic ambiguities

When the cake was dropped flat plants stuck to its underside

- Syntax A: *the cake was* **dropped** .... **flat plants** *stuck to its underside* (= modifying construction, [V [A N]])
- Syntax B: *the cake was* **dropped flat** .... **plants** *stuck to its underside* (= resultative construction, [[V A] N])

⇒ Corpus study:
Syntax A (=modifying) is far more likely than Syntax B (=resultative)

	[V A]	**[A N]**
ICE-GB/Brown	˜ 5/3%	˜ 67/88.5%

Experimental design: placement of phrase boundaries

- Difference in syntax comes with difference in the placement of an intonational phrase boundary

V **%** A N

or

V A **%** N

- Expect **V%AN** to occur more often (if speakers are given a choice)

→ corresponding syntactic structure is more frequent

Experimental design: lexical frequencies I

In order to determine:

1. effects of frequency **on syntactic choice**, the relevant syntactic sequence had to have four combinations:

Verb	Adj.	Noun	Shortcut
$V_{frequent}$	Adj.	$N_{frequent}$	ff
$V_{frequent}$	Adj.	$N_{infrequent}$	fi
$V_{infrequent}$	Adj.	$N_{frequent}$	if
$V_{infrequent}$	Adj.	$N_{infrequent}$	ii

2. effects of frequency **on boundary strength**, the four combinations above had to be comparable:
   - in the rhyme/coda of the verb
   - in the onset of the noun
   - in the onset and the rhyme/coda of the adjective

   → known to show the largest durational effects of boundary strength

   But: had to allow for reliable measurements at the same time

Experimental design: lexical frequencies II

Estimation of lexical frequencies via WebCelex:

	Verbs	**Nouns**
frequent	> 2000	> 3000
infrequent	< 200	< 100

Table: Raw number thresholds for lexical (in)frequencies

$\rightarrow$ Matching of verbs/nouns with respect to the form

ff: *When the cake was d**ropped** flat **p**lants stuck to its underside*

fi: *When the cake was d**ropped** flat **p**lanks stuck to its underside*

if: *When the grass was c**ropped** flat **p**lants were able to grow again*

ii: *When the grass was c**ropped** flat **p**lanks were laid across the lawn*

Bigram frequencies

Determined **bigram frequencies** of Verb-Adj (V-A) and Adj-Noun (A-N) combination and their **ratio**: V-A/A-N

Problem: No corpus large enough to determine frequencies of infrequent combinations.

$\rightarrow$ Google

$\rightarrow$ 'Noisy', therefore just approximations

$\rightarrow$ Great variance

$\Rightarrow$ Divided data into abstract categories:

	low	**med** (buffer)	**high**
	$< 40\%$	$40\% - < 60\%$	$>= 60\%$
V_A	< 13900	< 314000	$>= 314000$
A_N	< 3180	< 108000	$>= 108000$

Table: Abstract representation of raw bigram frequencies

Data gathering

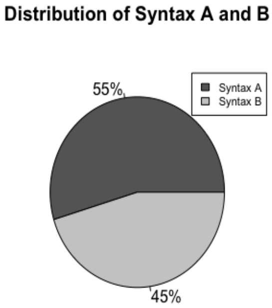
- Data presentation:
  - → without commas (syntactic boundary placed according to choice)
  - → several repetitions; only discuss first repetition here (58 sentences/speaker)
- Subjects: 23 participants
  (students at the University of Edinburgh, Ø=23,4 years, 14 females)
- Recordings: sound-treated studio at the University of Edinburgh with a high quality microphone

Frequency and syntactic choice: results I

Annotation of syntactic choice:
1 annotator (100%), 1 annotator (40%) – 100% agreement

Here: 23 speakers, repetition 1 → total of 1314 instances

Syntax A → V%AN
Syntax B → VA%N



... surprising given the results from the corpora ....

Frequency and syntactic choice: results II

For the choice of syntax, the following factors were relevant:

- **Syntax A** (frequent syntax, V%AN) more likely with
  - highly frequent nouns ($p < 0.05$)
  - high A-N bigram frequency ($p < 0.001$)
- **Syntax B** (infrequent syntax, VA%N) more likely with
  - highly frequent verbs ($p < 0.001$)
  - high V-A bigram frequency ($p < 0.001$)
  - higher V-A in comparison to A-N bigram frequency ($p < 0.001$)

Durational measurements: preparation

Strict selection:

- Only speakers that generally had a high consistency across repetitions (1 sentence - 1 choice - in both repetitions)

→ 10 speakers

- Only quadruplets that had the **same syntactic choice across both repetitions**

→ can measure frequency impact on duration – and later compare it to repetition 2

- Today: Discuss only repetition 1

Annotated sentences	
Syntax A	*Syntax B*
124	54

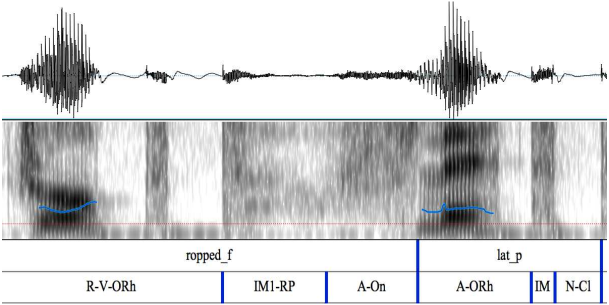
Durational measurements: annotation

- Raw material, e.g.

V-A	A-N	example
ropped_f	lat_t	*dropped flat plants*
k_f	ree_p	*walk free people*

→ Problematic, a lot of segmental variation

- Abstract annotation scheme, three intervals per sequence (six in total)



Durational measurements: annotation

Verb end		Adjective start		Adjective end		Noun start	
V-Rh	rhyme	A-On	onset	A-Rh	rhyme	N-On	onset
V-Co	coda	A-Cl	closure	A-Co/Co1/Co2	coda/coda part 1/2	N-Cl	closure
V-ORh	with part of onset			A-ORh	with part of onset		
R-V-...	with onset release			A-Nu	nucleus, not coda		
				R-A-...	with onset release		

Intermediate (IM1 **and** IM2)		**Comment:**
...-R	release	*Might include aspiration!*
...-P	pause	*Missing pause (P) is only indicated if syntax requires it*
...-RP	release and pause	*Both -P/-RP are only indicated if there is no closure following*
		If no R/P is present and not expected, then leave out IM. Else use brackets ()

Supra-markers		**Comment:**
?	insecurity	*Insecurity in annotation, mostly at preceding or following border*
x_x	connection	*Connection across word boundaries - e.g., V-Rh_IM1_A-On*
( )	missing element	*For elements that should be there, but are not (mostly R and P)*
NA		*If a separation at word boundary in DurationSep (only!) is not possible*
rel	release	*Only on DurationSep level. Connected to other parts with +*
pause	pause	*Same as release*
glot	glottalization	*Same as release*
(breath)	non-expected release	*Same as release*

→ Allows for grouping of similar patterns to get more reliable measurements!

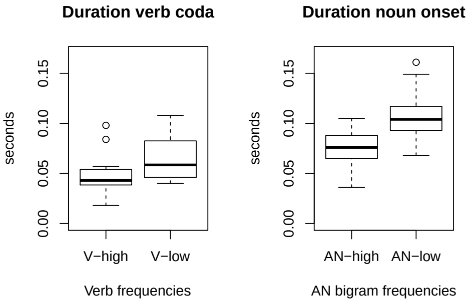
BUT: If there was no clear boundary, intervals were connected via an underscore (_)

→ particular item then not part of analysis - further reduction of data

Frequency and duration: some (significant) results I

Syntax A (frequent, V%AN):

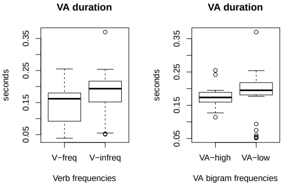
- When lexical frequency V is low: **increased verb coda interval duration** (β=0.015, SE=0.006, t=2.5, $p < 0.05$)
- When bigram frequency AN is low: **increased noun onset interval duration** (β=0.01, SE=0.004, t=2.3, $p < 0.05$)



Frequency and duration: some (significant) results II

Syntax B (infrequent, VA%N):

- When lexical frequency V is low or bigram frequency VA is low: **increase overall VA duration**
  (β=0.023, SE= 0.009, t=2.65, $p < 0.05$ and β=0.029, SE=0.01, t=2.96, $p < 0.01$)



→ Same effect is found with the verb coda (but not with the adjective onset)

- When VA bigram frequency higher than AN frequency:
  - **decrease of verb coda interval duration**
    (β =-0.029, SE= 0.007, t=-3.97, $p < 0.001$)
  - **increase of noun onset interval duration**
    (β=0.018, SE=0.005, t=3.27, $p < 0.01$)

Conclusion

All of these results are consistent with the SSRH:

→ inverse relationship between language redundancy (lexical frequencies, bigram frequencies, and their interaction) and durational measurements of the prosodic boundary-related intervals

Outlook:

- Compare repetitions
- Investigate F0
- Zoom in on bigram frequencies across boundaries
- ...

Thank you!

... questions, comments...?

EXTRAS

Information on corpora

	Brown corpus	**ICE-GB**
Released	1964	1998
Tagging	Part of Speech (POS)	Syntactic (Treebank)
Tokens	˜ 1 Million	˜ 1 Million
English	BE	AE
Texts	Across all genres	Edited English prose
Citation	(Francis and Kučera 1964)	(*ICE-GB corpus* 1998)

Table: Information on the ICE-GB and the Brown corpus

Results corpus study

Frequency determination:

	Verb-Adj		Adj-Noun
	main	copula	
ICE-GB corpus	1771	8781	21183
	10552		
In %	˜ **5%**	˜ 28%	˜ **67%**
	˜ 33%		
Brown corpus	1657	4562	47830
	10552		
In %	˜ **3%**	˜ 8,5%	˜ **88,5%**
	˜ 11,5%		

Table: Frequency of syntactic combinations in the ICE-GB and the Brown corpus

Conclusion:
Syntax A (=modifying) is more likely than Syntax B (=resultative)

Experimental design: Material

Examples with four combinations:

freq Verb	infreq Verb	freq Nouns	infreq Nouns
dropped	cropped	plank	plant
buy	dye	paper	paisley
call	wall	door	dorm
made	shade	picture	pitcher
make	rake	field	fief
stayed	bayed	sister	sissy
play	slay	fish	fiend
shake	snake	boxes	bobbers
turned	churned	balls	baulks
wear	pare	farmers	farthings
works	lurks	markets	marshals
walk	stalk	people	peafowls

References I

Aylett, Matthew. 2000. *Stochastic suprasegmentals: Relationships between redundancy, prosodic structure and care of articulation in spontaneous speech*. Ph. D.thesis, University of Edinburgh.

Aylett, Matthew and Turk, Alice. 2004. The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence, and Duration in Spontaneous Speech. *Language and Speech* 47(1), 31–56.

Aylett, Matthew and Turk, Alice. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *Journal of the Acoustical Society of America* 119(5), 3048–3058.

Bell, Alan, Brenier, Jason M., Gregory, Michelle, Girand, Cynthia and Jurafsky, Dan. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60(1), 92–111.

Francis, W. Nelson and Kučera, Henry. 1964. A Standard Corpus of Present-Day Edited American English, for use with Digital Computers. Technical Report, Department of Linguistics, Brown University, Providence, Rhode Island, US.

Gahl, Susanne and Garnsey, Susan. 2004. Knowledgef of grammar, knowledge of usage. Syntactic probabilities affect pronunciation variation. *Language* 80, 748–775.

ICE-GB corpus. 1998. ˙Online resource: http://ice-corpora.net/ice/icegb.htm .

Jurafsky, Dan, Bell, Alan, Gregory, Michelle and Raymond, William. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In Joan Bybee and Paul Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, pages 229–254, Amsterdam: John Benjamins.

Pluymaekers, Mark, Ernestus, Mirjam and Baayen, R. Harald. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118, 2561–2569.

Watson, Duane G., Breen, Mara and Gibson, Edward. 2006. The role of syntactic obligatoriness in the production of intonational boundaries. *Journal of Experimental Psychology: Learning, Memory and Cognition* 32, 1045–1056.