Why do we Accent Words?
The Processing of Focus and Prosodic Structure

Sasha Calhoun
University of Edinburgh
Sasha.Calhoun@ed.ac.uk

Experimental and Theoretical Advances in Prosody, Cornell, April 2008

Overview

- Why do we accent words?
  - focus and “informativeness”
  - metrical prosodic structure
- The relationship between prosody and focus
  - probabilistic constraints
- Accent prediction in the Switchboard corpus
- Prosody and processing
  - how, why and when do prosody and information affect language processing?

Focus and "Informativeness"

- Accents can mark **focus**
  - introduces a presupposition of a contextually appropriate set of alternatives to that element, to be resolved in the discourse
  - affects the interpretation (especially in terms of truth-conditional implications) of the utterance in relation to the discourse model (Rooth, 1992)

A: What would be the first thing you'd cut [in the budget]? Defence?
B: I would cut [the PRISON systems]F

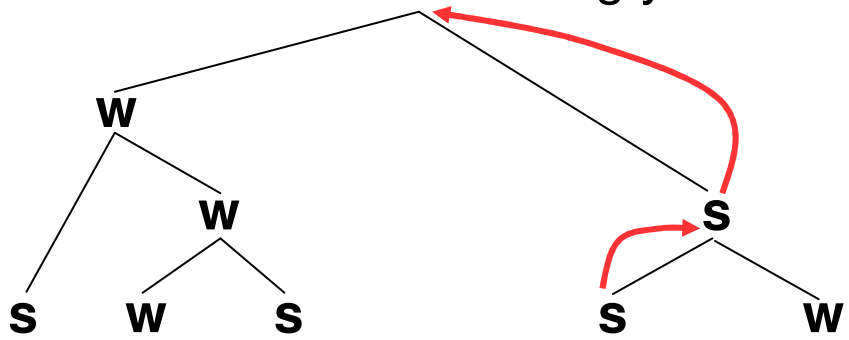
- *And* accents are more likely if word is '**informative**' or 'newsworthy'
  - Part-of-speech, predictability, "unexpectedness" and lexical '**accentability**' all strong predictors of accent status

(Bolinger 1972, Pan et al. 2002, Nenkova et al. 2007)

Relationship to Metrical Prosodic Structure

- Accents are also manifestations of metrical prominence structure

A: What would be the first thing you'd cut [in the budget]? Defence?



B: *I* would CUT [the PRISON SYSTEMS]F

- Words map onto a rhythmical, right-branching metrical structure (Liberman 1975)
- Foci align with *nuclear* accents
  - addresses problem of 'non-focal' accents

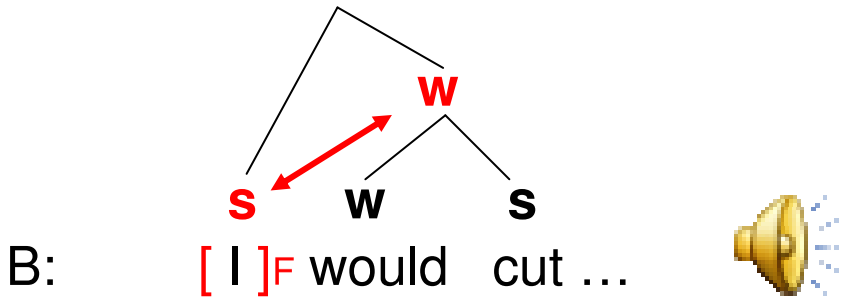
(Ladd 1996, Truckenbrodt 1995, Büring 2006)

My Model

- Mapping of words onto metrical prosodic structure is **probabilistic**
  - focus is a strong constraint on this mapping
- The formation of metrical structure and the alignment with words is also governed by other constraints, e.g.
  - maintenance of rhythmic alternation
  - physiological pressure on phrase length
  - right-branching bias in metrical structure
  - overall emphasis
  - 'accentability' of word given part-of-speech, argument status (subject versus object), predictability
  - alignment of syntactic and prosodic phrases

Probability-based Expectation

- The likelihood of any word being accented (or nuclear accented) depends on (at least) its place in prosodic structure, its syntactic properties and its focal status
- Where the prosodic structure associated with a word sequence is not expected, we get a marked reading



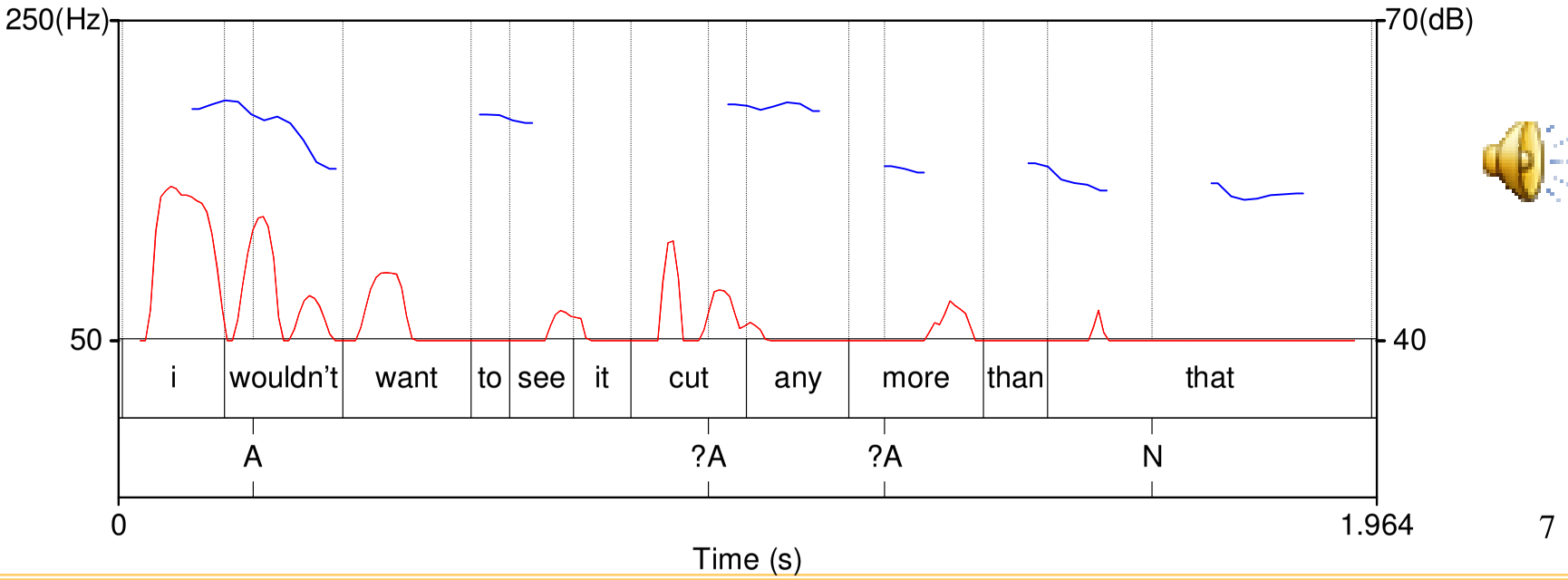
- Foci can be marked by non-nuclear accents when the accented word is more prominent than expected

- Where the nuclear accent falls where expected, we do not need as much phonetic prominence to mark focus
  - this explains why focal accents early in a phrase are generally much higher than focal accents late in a phrase

(Rump & Collier 1996, Xu & Xu 2005)

- The phonetic prominence and tonal shape of a 'default' nuclear accent can vary to convey other information

As for the defence budget, they're already cutting it back what 25%…



Accent Prediction in Switchboard

- Corpus of spontaneous telephone conversations annotated for syntax, focus, prosody and information status (Calhoun et al. 2005)
- Only included words in fluent phrases, annotated for focus status
  - approx 9000 words, 2 ½ hours
- Multinomial logistic regression modelling
- Used models to classify each word as having a *nuclear accent, non-nuclear accent* or *no accent*
- Primarily interested in importance of different feature sets

Results (%)

	None	Accent	Nuclear	Overall
Baseline	100	0	0	47.2
Prosodic	82.9	29.0	71.9	61.5
Syntactic/Semantic	77.1	13.2	69.5	57.4
Pros+Syn/Sem+WdPr	85.4	48.4	77.2	67.6

- Nuclear accents can be predicted much more reliably than non-nuclear accents

	None	Accent	Nuclear	Overall
Baseline	100	0	0	47.2
Prosodic	82.9	29.0	71.9	61.5
Syntactic/Semantic	77.1	13.2	69.5	57.4
Pros+Syn/Sem+WdPr	85.4	48.4	77.2	67.6

- Prosodic features best distinguish non-nuclear accents

	None	Accent	Nuclear	Overall
Baseline	100	0	0	47.2
Prosodic	82.9	29.0	71.9	61.5
Syntactic/Semantic	77.1	13.2	69.5	57.4
Pros+Syn/Sem+WdPr	85.4	48.4	77.2	67.6

	Feature	% Difference	
		Non-Nuclear	***Nuclear***
Increase	focus	+ 14.9	+ 25.3
	focus & verb	-	+ 13.8
	num syllables in word	+ 17.4	+ 19.3
	break follows	-	+ 18.9
	words since last accent	+ 5.4	+ 5.8
	relative duration	+ 14.8	+ 18.0
	mean intensity	+ 6.5	+ 8.8
	mean pitch	+ 4.6	+ 7.3
Decrease	pronoun	– 8.8	-
	verb	– 5.3	-
	relative position in phrase	– 18.8	-
	accents in phrase so far	– 14.1	– 21.7
	phrase mean pitch	– 2.9	– 5.2

	Feature	% Difference	
		Non-Nuclear	*Nuclear*
Increase	focus	+ 14.9	+ 25.3
	focus & verb	-	+ 13.8
	num syllables in word	+ 17.4	+ 19.3
	break follows	-	+ 18.9
	words since last accent	+ 5.4	+ 5.8
	relative duration	+ 14.8	+ 18.0
	mean intensity	+ 6.5	+ 8.8
	mean pitch	+ 4.6	+ 7.3
Decrease	pronoun	– 8.8	-
	verb	– 5.3	-
	relative position in phrase	– 18.8	-
	accents in phrase so far	– 14.1	– 21.7
	phrase mean pitch	– 2.9	– 5.2

	Feature	% Difference	
		Non-Nuclear	***Nuclear***
Increase	focus	+ 14.9	+ 25.3
	focus & verb	-	+ 13.8
	num syllables in word	+ 17.4	+ 19.3
	break follows	-	+ 18.9
	words since last accent	+ 5.4	+ 5.8
	relative duration	+ 14.8	+ 18.0
	mean intensity	+ 6.5	+ 8.8
	mean pitch	+ 4.6	+ 7.3
Decrease	pronoun	– 8.8	-
	verb	– 5.3	-
	relative position in phrase	– 18.8	-
	accents in phrase so far	– 14.1	– 21.7
	phrase mean pitch	– 2.9	– 5.2

Summary

- Both nuclear and non-nuclear accents best predicted by prosodic structure features
- Clear difference in distribution of non-nuclear and nuclear accents:
  - nuclear accents much more linked to focus
  - non-nuclear accents predicted by low-level features, e.g. part-of-speech
  - prediction of non-nuclear accents much worse
- 'High' and 'low'-level informational features have independent effects

Why do we accent words?

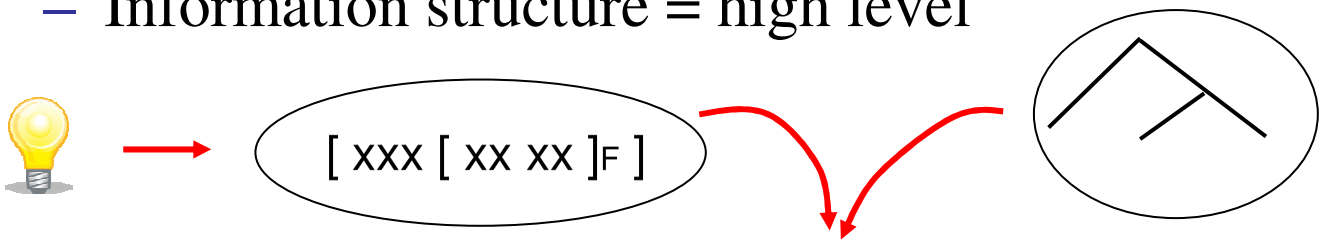
- Because prosodic structure tells us to…
  - but why should there be prosodic structure?
- Prosody acts to hold the different levels of linguistic representation of each utterance together

(Frazier et al. 2006)

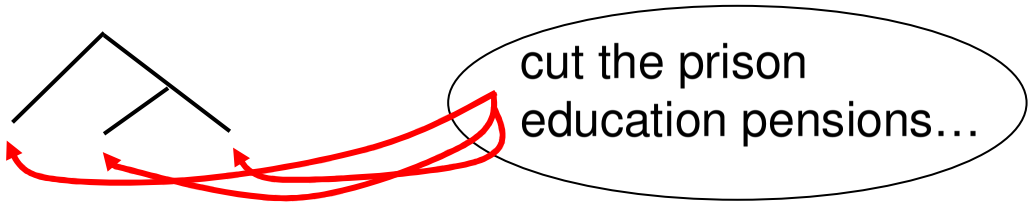
- Prosodic structure, i.e. phrasing *and* prominence, is good for this because it is **predictable**:
  - exploits our pattern recognition ability
- Having a ready frame in which to slot each information unit aids processing
- *and* we can exploit this predictability for extra effect

How does information affect prosody?

- We have seen two kinds of "information" effects:
  - Information structure: focus ↔ nuclear accent
  - 'Accentability': part-of-speech, givenness, predictability
- Act at different levels of utterance production:
  - Information structure = high level



  - 'Accentability' = low-level → linked to lexical retrieval



- Ambiguity arises when the two processes lead to similar outcomes

Conclusion

- Presented model of the relationship between prosodic structure and focus
  - focus acts as a strong constraint on the probabilistic mapping of words onto metrical structure
- Corpus Results: accent prediction in Switchboard
  - prosodic structure strong predictor of both nuclear and non-nuclear accents
  - nuclear accenting more strongly related to focus, non-nuclear accents to 'low-level' features like part-of-speech
- Prosody and processing
  - metrical prosodic structure binds utterance representations
  - focus part of high-level utterance/prosody planning; 'accentability' part of low-level planning

References

Bolinger, D. (1972). Accent is predictable (if you're a mind reader), *Language* 49, 633–644.

Büring, D. (2006), Focus projection and default prominence, in V. Molnár & S. Winkler, eds, *The Architecture of Focus*, Mouton De Gruyter, Berlin/New York.

Calhoun, S., Nissim, M., Steedman, M. & Brenier, J. (2005), A framework for annotating information structure in discourse, in 'Frontiers in Corpus Annotation II: Pie in the Sky', *ACL2005 Conference Workshop*, Ann Arbor, MI.

Frazier, L., Carlson, K. & Clifton Jr., C. (2006). Prosodic phrasing is central to language comprehension, *Trends in Cognitive Science* 10, 6, 244-249.

Ladd, D. R. (1996), *Intonational Phonology*, Cambridge University Press, Cambridge, UK.

Liberman, M. (1975), *The intonational system of English*, PhD thesis, MIT Linguistics, Cambridge, MA.

Nenkova, A., J. Brenier, A. Kothari, S. Calhoun, L. Whitton, D. Beaver & D. Jurafsky (2007). To memorize or to predict: Prominence labeling in conversational speech. In *HLT:NAACL*, Rochester, NY.

Pan, S., McKeown, K. & Hirschberg, J. (2002), Exploring features from natural language generation for prosody modeling, *Computer Speech and Language* 16, 457–490.

Rooth, M. (1992). A theory of focus intepretation, *Natural Language Semantics* 1, 75–116.

Rump, H. & Collier, R. (1996), Focus conditions and the prominence of pitch-accented syllables, *Language and Speech* 39, 1–17.

Truckenbrodt, H. (1995), *Phonological Phrases: Their Relation to Syntax, Focus and Prominence*, PhD thesis, MIT.

Xu, Y. & Xu, C. (2005), Phonetic realization of focus in English declarative intonation, *Journal of Phonetics* 33, 159–197.