

Prosodic features of stance strength and polarity

Valerie Freeman
Indiana University

LSA Annual Meeting
Washington, DC
January 9, 2016



Project

- ATAROS
 - Automatic Tagging and Recognition of Stance
 - Collaboration with phoneticians, computational linguists, signal-processing engineers
 - Corpus hosted at the University of Washington
 - Seeks automatically-extractable acoustic cues to stance
 - Also Marvel god of video games →



Terms

- Stance
 - Speaker's attitudes, opinions, feelings, judgments about topic of discussion (Biber et al. 1999; Conrad & Biber 2000)
 - Related: evaluation, attitude, sentiment, subjectivity
 - Stance-taking: Activity of expressing stance (Haddington 2004)
 - Essential to collaboration, negotiation, decision-making

Related Work

- **Conversation Analysis & Discourse Analysis**
 - Qualitative, often small amounts of data
 - (e.g., Biber & Finegan 1989, Conrad & Biber 2000, Du Bois 2007, Englebretson 2007, Haddington 2004, Hunston & Thompson 2000, Jaffe 2009, Ogden 2006)
- **Computational Linguistics/Speech Recognition**
 - Often relies on text or lexical features, but much more information is available in the speech signal
 - (e.g., Murray & Carenini 2009, Hillard et al. 2003, Somasundaran et al. 2006, Wilson 2008, Wilson & Raaijmakers 2008, Raaijmakers et al. 2008)

ATAROS Corpus

- High-quality audio
- 34 dyads from Pacific Northwest
 - Strangers matched by age
- 5 stance-dense collaborative tasks
- Transcribed, time-aligned to audio
- Annotated for stance strength, polarity, type
- Available to other researchers

Tasks

| | Neutral first-mentions | Increasing involvement |
|-----------------|---------------------------|---------------------------|
| Store items | Map | Inventory Survival |
| Budget items | Category | Budget |

Inventory Task

- Scenario: You're co-managers of a new superstore in charge of arranging inventory
- Decide together where to place each target item on a felt wall map
- Low involvement, weak opinions, agreement

Inventory Task

- W- We should-
- So, fridge-
- We should- make a- a- a decision where beverages should go, anyway. So, it doesn't-
- Yeah.
- I don't think it's a big... huge decision to s-
- We could do b- beverages like here.
- Sure.
- Maybe.
- Perfect.



Budget Task

- Scenario: You're on the county budget committee, and it's time to make cuts
- Decide together which expenses to cut from each department
- High involvement, stronger opinions, more persuasion, reasoning, negotiation, personal experience as support

Budget Task

- {breath} Alright. .. Wh- Poetry books .. or cooking classes?
- No, if you're gonna leave in football, we need poetry.
- Oh we're not g- Oh - oh, I'm willing to take out - {breath}
- Oh, football equipment?
- Yeah.
- Oh.
- So if we take out the juice machines and football, we've done it.
- Okay.



Transcription & Annotation

- Manual orthographic transcription in Praat (Boersma & Weenink 2013)
- Forced-alignment w/ P2FA (Yuan & Liberman 2008)
 - Aligns word and phone boundaries with audio
- Manual stance annotation
 - 2-3 annotators label stance strength and polarity of each “spurt” (utterance between >500ms silences) via content analysis (modified from Freeman 2014)

Stance Strength

- Each spurt marked for stance strength:
 0. None: reading, backchannels, facts
 1. Weak: cursory agreement, suggest solution, solicit opinion, mild opinion/reasoning
 - “What do you think?” / “Sure.”
 2. Moderate: stronger agreement, opinion, reasons; disagreement, alternate solutions
 - “Let’s do this instead.”
 3. Strong: very strong versions of above
 - “What?! Screw that!”

Polarity

- Spurts with stance also marked for polarity:
 - + Positive: Agreement, encouragement
 - Negative: Disagreement, hedging, questioning other's opinion
 - ∅ Neutral/neither: offer, solicit opinion

Hypothesis & Measures

- Measurable cues to stance strength and polarity are present in the acoustic signal
 - Same words, different messages...
- Automatically-extracted measures:
 - Pitch, intensity at vowel midpoint & every decile
 - Z-score normalized within speaker
 - Vowel duration
 - Z-score normalized within speaker & vowel quality

Data Set

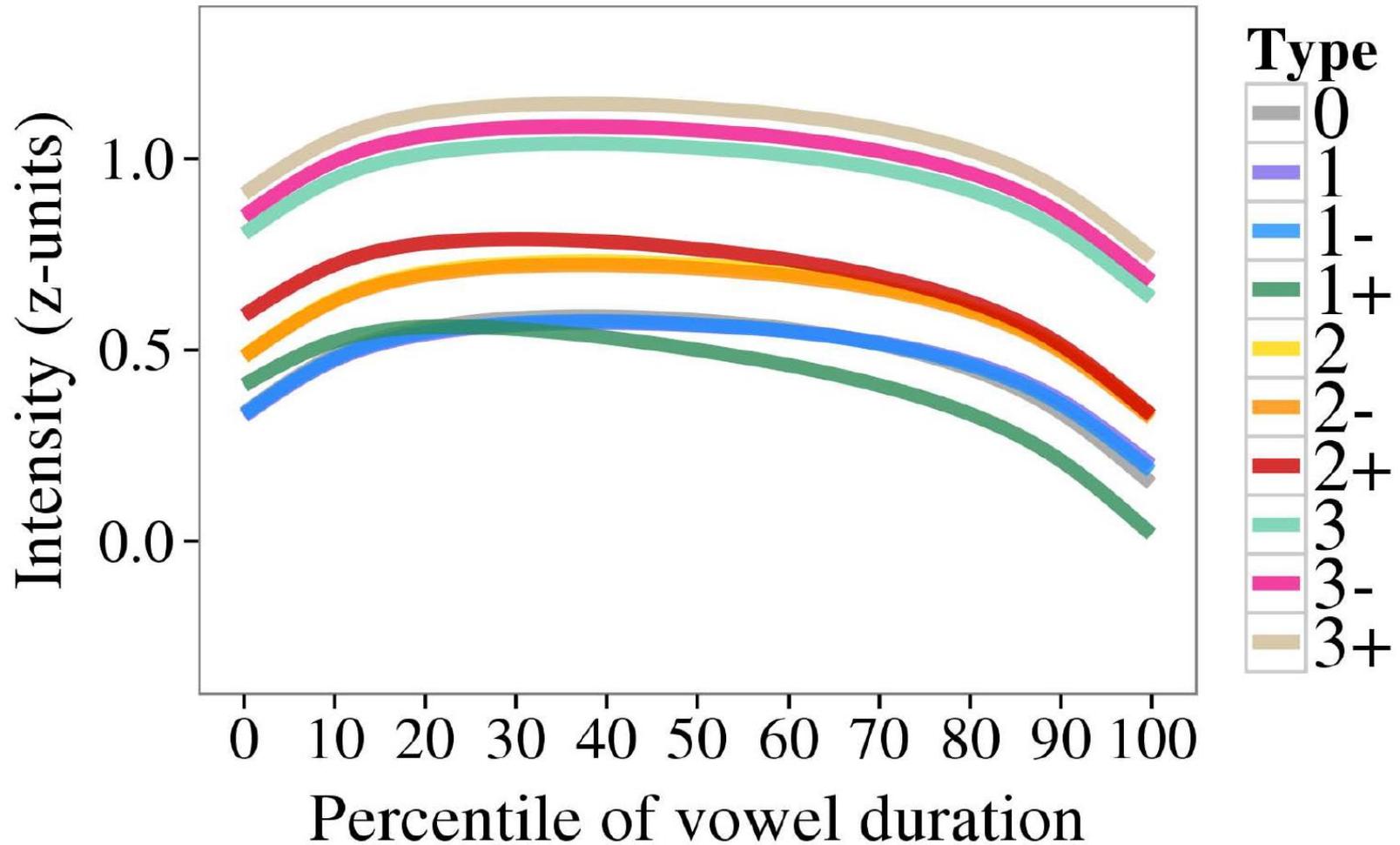
- 20 dyads
 - Dyads: 7 FF, 3 MM, 10 mixed-sex
 - Speakers: 24 F, 16 M (half under age 35)
- Inventory & Budget task data combined
- 32,000 stressed vowels from content words

Intensity

- Increases with stance strength ($p < 0.001$)
 - Except: 1+ lowest
- Low 1+ brings positive polarity average down
- Combined labels cluster by strength

| | | | |
|---------|-------|----|----|
| highest | 3 | 3+ | 3- |
| | <hr/> | | |
| | 2 | 2- | |
| | <hr/> | | |
| | 0 | 1 | 1- |
| | <hr/> | | |
| lowest | 1+ | | |

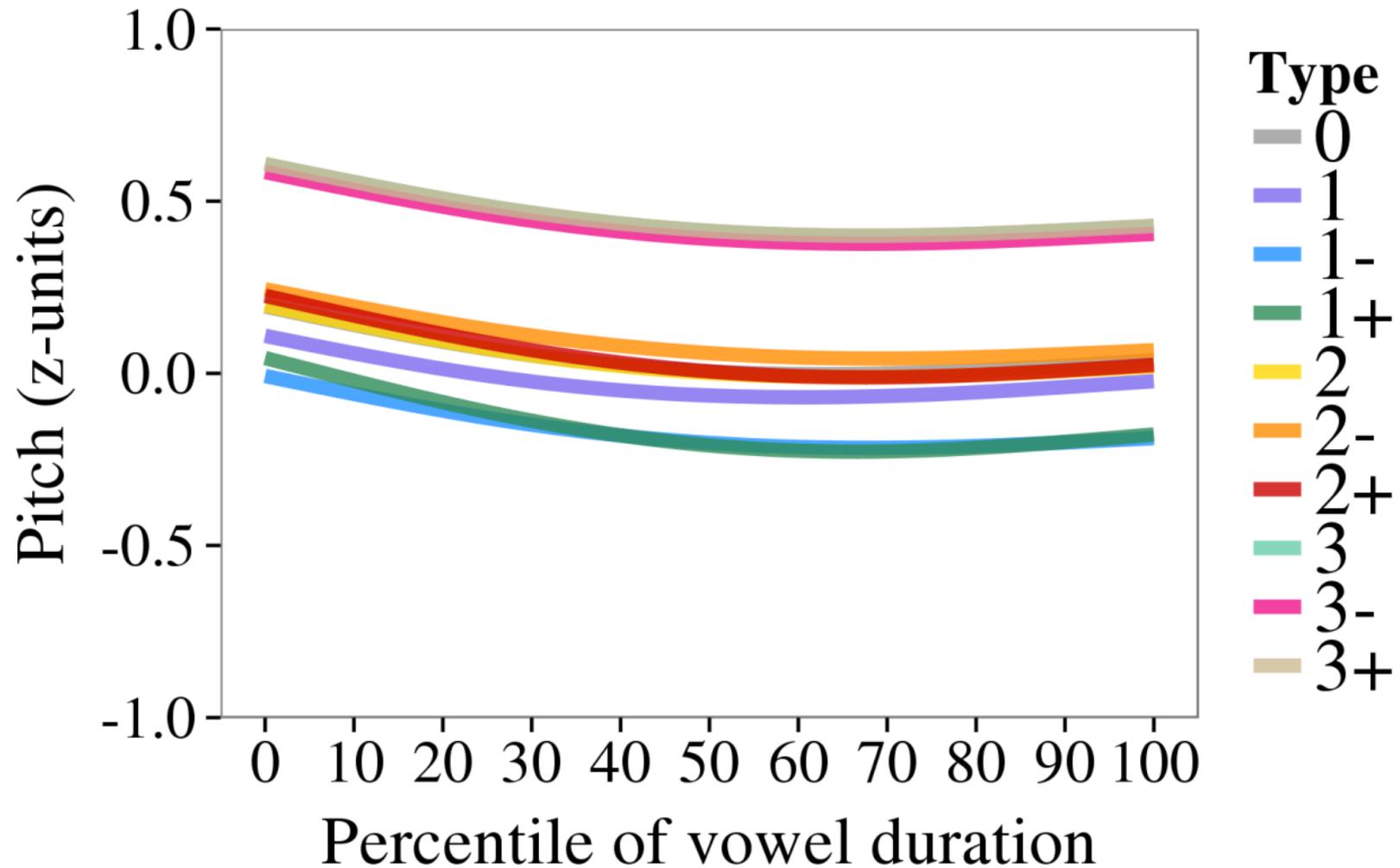
Intensity Contours



Pitch

- Increases with stance strength ($p < 0.001$)
 - Low 1+ brings positive polarity average down
 - Most strength/polarity combined labels don't differ from immediate neighbors
- | | |
|---------|-----------|
| highest | 3+ |
| | 3 |
| | <u>3-</u> |
| | 2 / 0 |
| | 2+ |
| | 1 |
| | 2- |
| | 1- |
| lowest | 1+ |

Pitch Contours



Vowel Duration

- Positive polarity longer ($p < 0.001$)
 - Neg/neutral don't differ
 - 1- differs only from 1+
 - 3+ doesn't differ from any
- Decreases with stance strength ($p < 0.001$)
 - Except: strong 3/3+/3- too variable, overlap all

| | | | | | |
|-------|---|----|----|----|----|
| long | 0 | 1+ | 2+ | 3 | 3+ |
| short | 1 | 1- | 2 | 2- | 3- |

Conclusion

- Measurable prosodic cues to stance strength and polarity are present in the acoustic signal

| | Pitch | Intensity | Duration |
|----------|--------------------------------|--------------------------------|--------------------------|
| Strength | Increases with strength levels | Increases with strength levels | |
| Polarity | | | Positive = longer vowels |

Conclusion

- Measurable prosodic cues to stance strength and polarity are present in the acoustic signal
 - Pitch and intensity increase with stance strength
 - Positive stances have longer stressed vowels (are said more slowly)
- Future work
 - Prosodic contours/tunes over speech acts
 - Social variables
 - Perception (ongoing at UW)

References

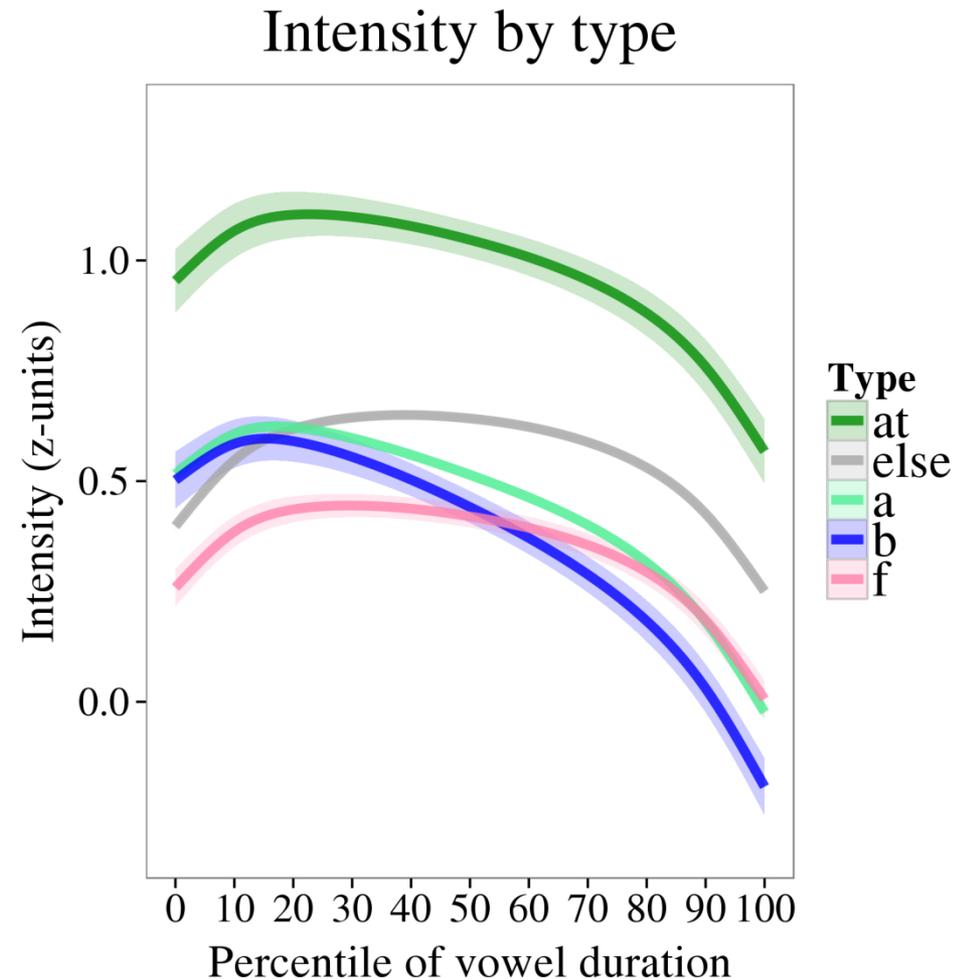
- Biber, D. & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text - Interdisciplinary Journal for the Study of Discourse*, 9(1):93-124.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. (1999). *Longman grammar of spoken and written English*.
- Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer, v. 5.3.
- Conrad, S. & Biber, D. (2000). Adverbial marking of stance in speech and writing. S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 56-73.
- Du Bois, J. (2007). The stance triangle. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139-184.
- Englebretson, R. (2007). Stancetaking in discourse: An introduction. *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 1-26.
- Freeman, V. (2015). *The phonetics of stance-taking*. Doctoral dissertation. Seattle: University of Washington Press.
- Freeman, V. (2014). Hyperarticulation as a signal of stance. *Journal of Phonetics*, 45, 1-11.
- Haddington, P. (2004). Stance taking in news interviews. *SKY Journal of Linguistics*, 17:101-142.
- Hillard, D., Ostendorf, M. & Shriberg, E. (2003). Detection of agreement vs. disagreement in meetings: Training with unlabeled data. *Proceedings of HLT-NAACL*.
- Hunston, S. & Thompson, G. (2000). Evaluation: An introduction. *Evaluation in text: Authorial stance and the construction of discourse*, 1-27.
- Jaffe, A. (2009). *Stance: Sociolinguistic Perspectives*.
- Murray, G. & Carenini, G. (2009). Detecting subjectivity in multiparty speech. *Proceedings of Interspeech*.
- Ogden, R. (2006). Phonetics and social action in agreements and disagreements. *J. of Pragmatics*, 38(10):1752-1775.
- Raaijmakers, S., Truong, K. & Wilson, T. (2008). Multimodal subjectivity analysis of multiparty conversation. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*.
- Somasundaran, S., Wiebe, J., Hoffmann, P. & Litman, D. (2006). Manual annotation of opinion categories in meetings. *Proceedings of Coling/ACL*.
- Wilson, T. (2008). Annotating subjective content in meetings. *Proceedings of Language Resources and Evaluation*.
- Wilson, T. & Raaijmakers, S. (2008). Comparing word, character, and phoneme n-grams for subjective utterance recognition. *Proceedings of Interspeech*.
- Yuan, J. & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Proceedings of Acoustics '08*.

Thanks

- Support: NSF IIS 1351034; NIH R01 DC60014; UW Excellence in Linguistic Research Graduate Award
- The ATAROS team (ataros@uw.edu):
 - PIs: Gina-Anne Levow, Richard Wright, Mari Ostendorf
 - Co-RAs: Yi Luan, Julian Chan, Trang Tran, Alena Hrynkevich, Victoria Zayats, Maria Antoniak, Sam Tisdale, Liz McCullough
 - Annotators: Heather Morrison, Lauren Fox, Nicole Chartier, Marina Oganyan, Max Carey, Andrew Livingston, Phoebe Parsons, Griffin Taylor
- Info/corpus access: depts.washington.edu/phonlab/projects.htm
 - This work, fully reported in: Freeman (2015) dissertation
 - My contact: vdfreema@iu.edu

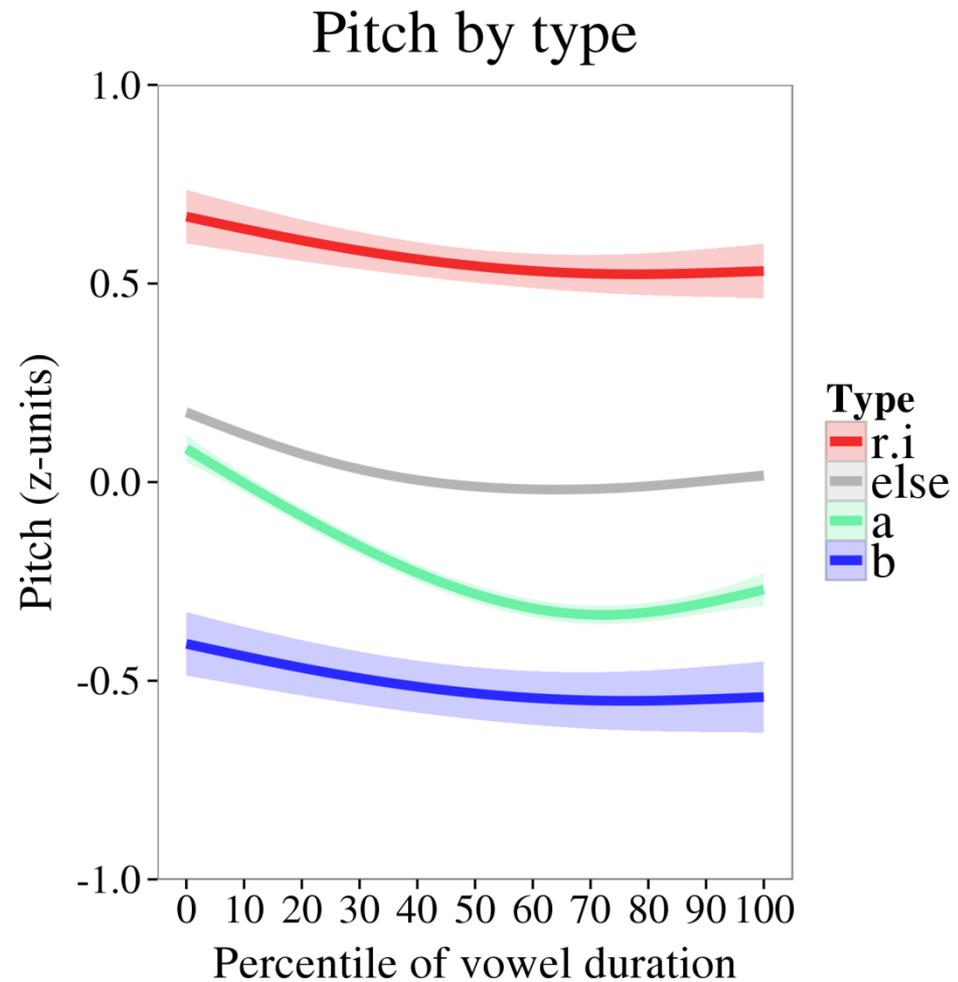
Intensity Contours by Type

- Clusters:
 - Rapport-building agreement (at) very high
 - Agreement (a) dropping
 - Backchannels (b) low dropping
 - Softening (f) low



Pitch Contours by Type

- Clusters:
 - Reluctance, strong intonation (r, i) high
 - Agreement (a) mod-low dipping
 - Backchannels (b) low



Results Summary

| Stance feature/type | | Pitch | Intensity | Duration |
|---------------------|-------------------------|--------------------------------|--------------------------------|-----------------|
| | Strength | increases with strength levels | increases with strength levels | – |
| | Polarity | – | – | positive longer |
| r; i | reluctance; intonation | very high | – | long |
| at | agreement+rapport | – | very high | very long |
| a,1+ | weak-positive agreement | low-dipping | dropping | long |
| b | backchannels | very low | low-dropping | long |
| f | softening/hesitation | – | low | – |