Modeling laryngeal acoustics in Early Korean Infant-Directed Speech: VOT and F0

Chandan Narayan
University of Toronto Scarborough
Tae-Jin Yoon
McMaster University
Derek Denis
University of Toronto

Acoustics Week in Canada, Quebec City 2011

Introduction

- Korean exhibits a three-way laryngeal contrast in oral stops and affricates: plain (lenis), tense (fortis), aspirated
- Voice onset time (VOT), post-release fundamental frequency (f0) (Lisker and Abramson 1964), and H1-H2 (Kang and Guion 2006) have been shown to systematically vary according to laryngeal type
- Evidence of a *sound change* with respect to the implementation of VOT as an acoustic correlate to the laryngeal distinction in speakers born after 1965 (Silva et al. 2004; Kang and Guion 2006)

Contrast Enhancement

- Research in infant-directed speech (IDS) suggests that certain phonetic contrasts are *enhanced* with respect to similar contrasts in adult-directed speech (ADS)
  - Vowel-space expansion (Liu et al. 2003, Werker et al. 2007)
  - Vowel length (Werker et al. 2007)
  - Tone contrast (Liu et al. 2007)

Contrast abatement in IDS

- Examples of enhancement in IDS are focused on duration dependent phonetic contrast (necessarily vocalic)
- Contrast *abatement* has been found in consonantal contrasts that require fine oral-laryngeal timing
- Voice onset time (VOT), the primary acoustic cue to voicing in many languages has been found to characterize English voiced and voiceless stops less accurately in IDS than in ADS (Narayan *forthcoming*, Baran et al. 1977)

Research questions

- How is a VOT implemented in a language that contrasts three laryngeal types?
- Do Korean-speaking mothers enhance or abate the three-way laryngeal contrast?
- How does early ID Korean differ from conversational Korean spoken by the same person?

Korean IDS/ADS Corpus

- Recordings retrieved from the Cross-Linguistic Corpus of Infant-Directed Speech (CCIDS) at the University of Toronto Scarborough
- Four mothers whose native language is Korean (aged 33-40) were recorded every month for 12 months (beginning when infants were 5 months)
  - Hour long recordings were made at mothers' homes while interacting with their infant (45 mins of IDS) and 15 min of adult-directed speech with a Korean speaking researcher

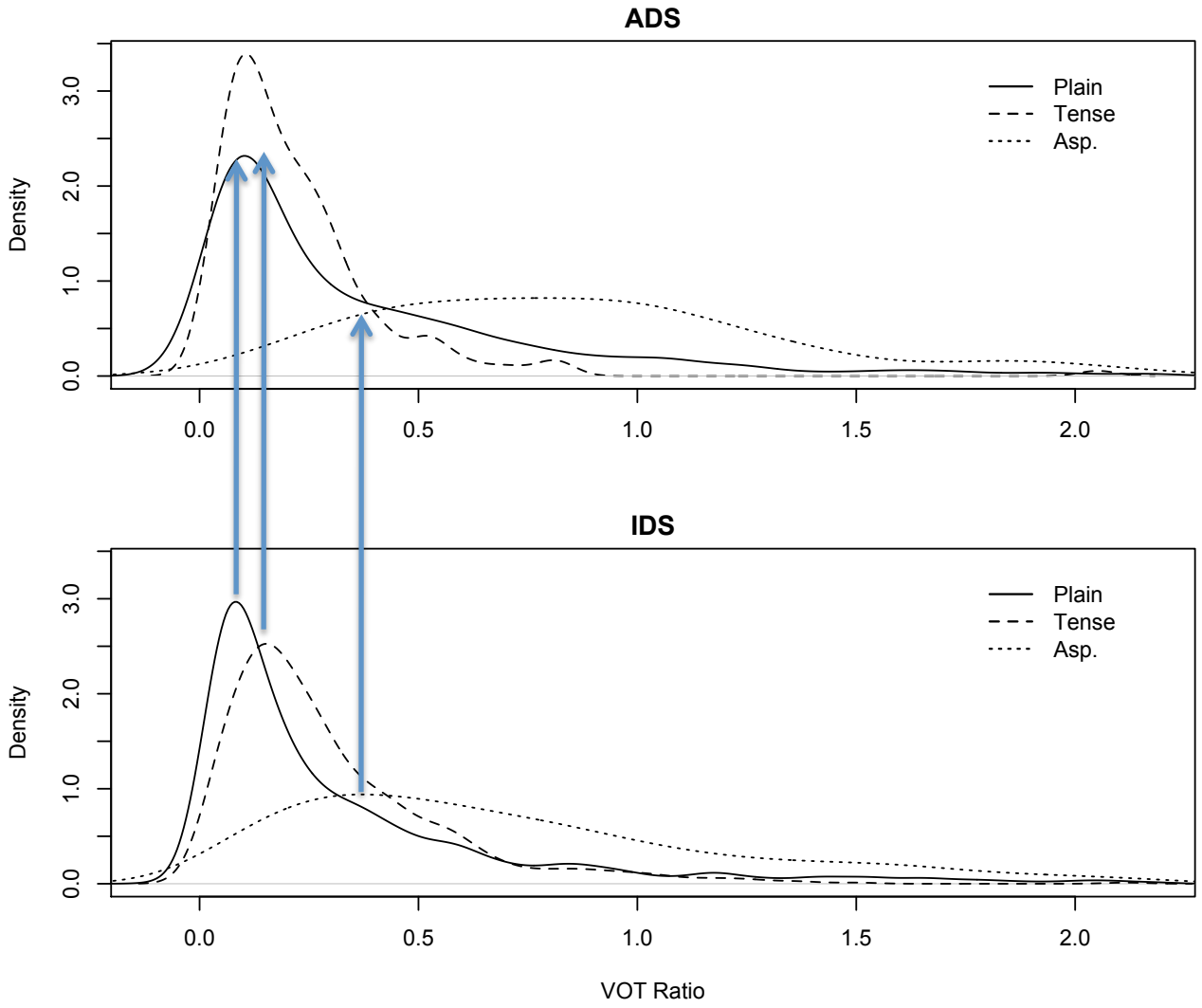
- CVs (oral obstruents) were tagged (in *Praat*) by a trained Korean-speaking phonetician
- VOT was measured as the duration between the first zero-crossing of the stop release (burst) and the onset of vocalization
- Peak F0 (Umeda 1981) in the first half of the post-obstruent vowel

Data

- Recordings of mothers speaking Korean with 5-6 month old infants
- Only word-initial CVs analyzed (~3000)
- VOT values were normalized to allow for comparison across the IDS and ADS registers

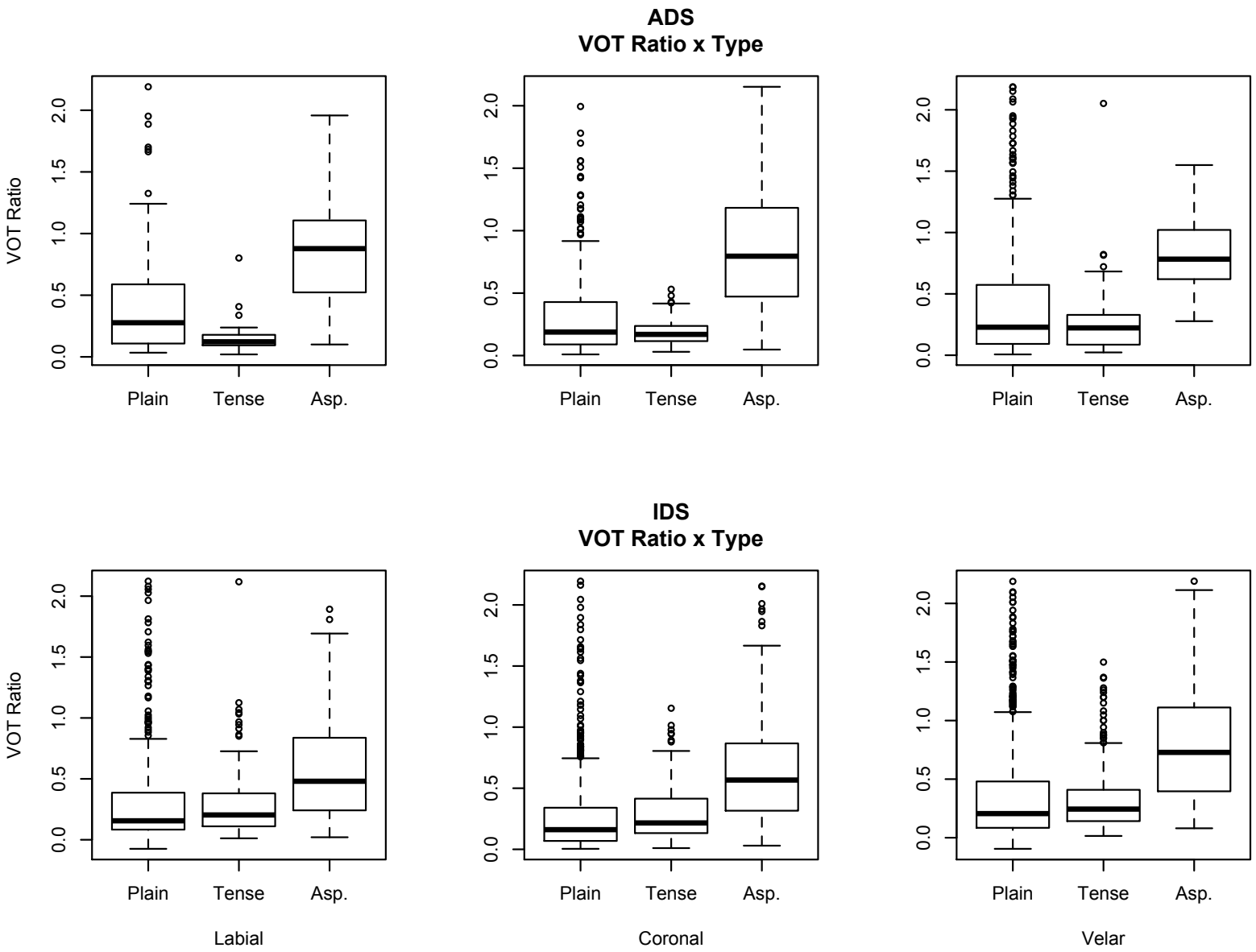
  VOT (*ms*)/Post-obstruent vowel duration = **VOT ratio**
- F0 (*Hz*) samples converted to *Mel* scale (Stevens and Volkmann 1940) and normalized by speaker using standard *z* calculation

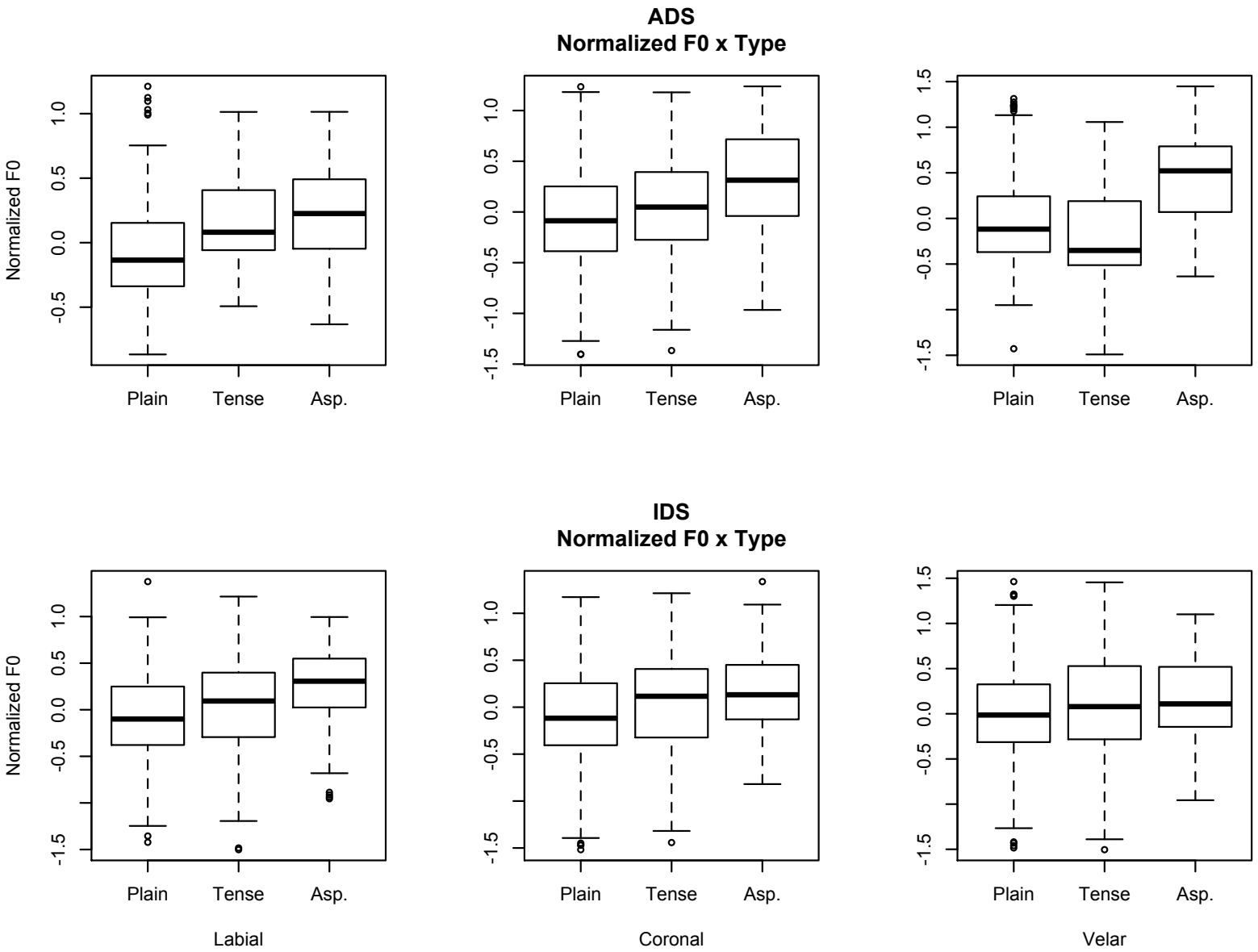
VOT probability Density

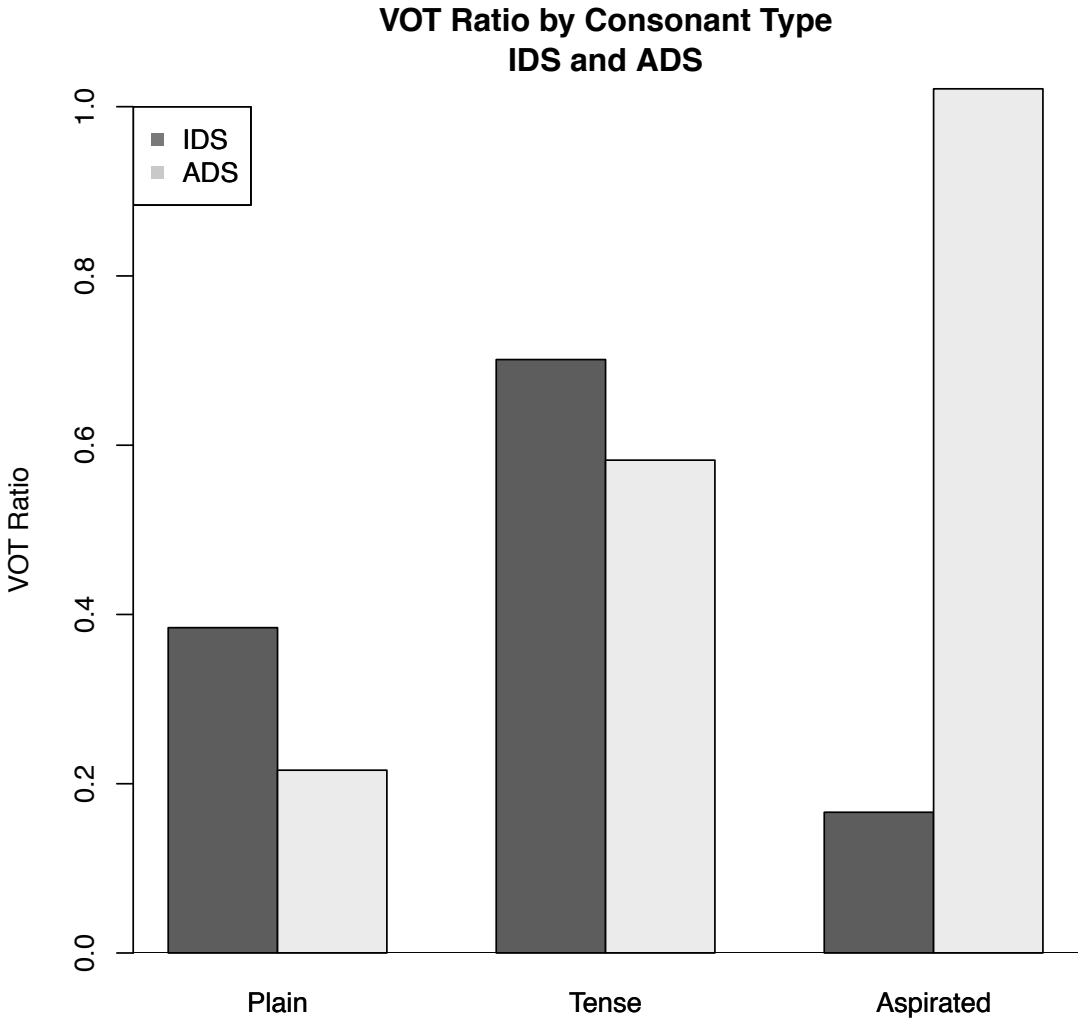


VOT and F0 in IDS vs. ADS registers

- VOT and F0 are implemented differently in IDS and ADS as suggested by and interaction between *LARYNGEAL TYPE* and *REGISTER* in a 2 (REGISTER: IDS, ADS) x 3 (TYPE: plain, tense, aspirated)
  - Interaction with place-of-articulation in the predictable direction
    - Velar stop have longer VOTs than bilabial and alveolar in both registers

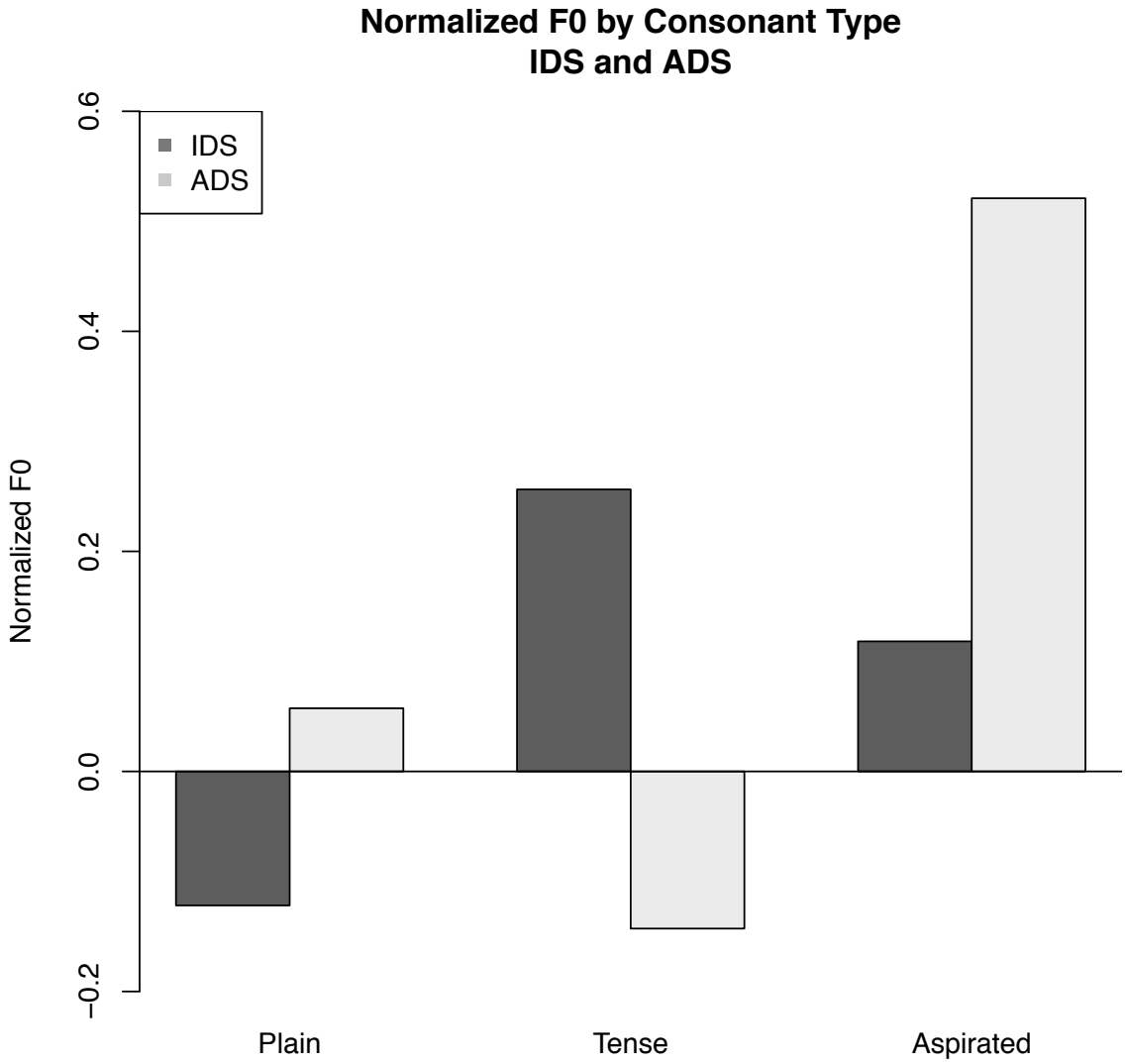

ADS
VOT Ratio x Type
IDS
VOT Ratio x Type
VOT Ratio
0.0
0.5
1.0
1.5
2.0
Plain
Tense
Asp.
Labial
Coronal
Velar


ADS
Normalized F0 x Type
IDS
Normalized F0 x Type
Normalized F0
Plain
Tense
Asp.
Labial
Coronal
Velar



- All within-register means show significant differences (by Wilcoxon tests)

- Within consonant type, only Aspirated stops in IDS vs. ADS showed a significant difference



- All IDS means are significantly different
- Within ADS, there is no difference between Plain and Tense F0

Predictive models of laryngeal type

- Multinomial logistic regression was used to model both individual and pooled, register specific predictions of consonant type
- Logistic regression model for more than two discrete outcomes (here Plain, Tense, and Aspirated consonant types) with VOT and F0 as predictors

IDS model (pooled)

	Estimate	Std.Error	t-value	Pr(>|t|)
Plain-Tense: VOT	-1.825698	0.240896	-7.5788	< 0.001
Plain-Tense: F0	0.024324	0.223528	0.1088	0.913
Plain-Tense: VOT x F0	1.183885	0.552063	2.1445	< 0.05
Plain-Aspirated: VOT	0.77169	0.173331	4.4521	< 0.001
Plain-Aspirated: F0	1.161061	0.322451	3.6007	< 0.001
Plain-Aspirated: VOT x F0	0.768566	0.418979	1.8344	0.07

ADS model (pooled)

	Estimate	Std.	Error	t-value
Plain-Tense: VOT	-5.524813	0.895983	-6.1662	< 0.001
Plain-Tense: F0	0.648645	0.58089	1.1166	0.2641483
Plain-Tense: VOT x F0	-0.809259	2.20011	-0.3678	0.7130026
Plain-Aspirated: VOT	1.312937	0.385073	3.4096	< 0.001
Plain-Aspirated: F0	0.994631	0.820943	1.2116	0.2256764
Plain-Aspirated: VOT x F0	2.41816	0.930464	2.5989	< 0.01

IDS models

- VOT ratio was predictive of PLAIN vs. TENSE for 1 of 4 speakers
- VOT ratio was predictive of PLAIN vs. ASP for all 4 speakers
- F0 was predictive of PLAIN vs. TENSE for 1 of 4 speakers
- F0 was predictive of PLAIN vs. ASP for 1 of four speakers for 3 of 4 speakers

Modeling conclusions

- Multinomial logit models of the laryngeal contrast in Korean IDS and ADS are characterized using VOT and F0
- In IDS, the Plain-Tense distinction is captured by covarying VOT and F0 values, as F0 alone does not contribute to the prediction

Conclusions

- Laryngeal contrasts in IDS Korean at 5mos is neither enhanced nor abated relative to AD Korean
- The phonetic contrasts are implemented differently in the two registers with the most notable modification being shorter VOT in Aspirated stops in IDS

Aspirated stops in IDS

- Korean, Swedish (Sundberg and Lacerda), and English (Baran et al. 1977, Narayan *forthcoming*) IDS show significant shortening of phonetically aspirated stops relative to ADS
- Why? → Higher overall F0(?)
- There is some research suggesting that raising F0 sacrifices oral-laryngeal timing (McCrea and Morris 2005)
- High pitch → short VOT potentially has consequences for the developing phonetic space in infancy

Next steps

- The CCIDS contains longitudinal recordings of Korean IDS
  - How does the ID Korean laryngeal contrast change (with respect to VOT and F0) as the infant gets older?
  - At what point in development does the contrast "look" like the contrast as implemented in ADS?
  - How does voice quality (H1-H2) contribute to a prediction of laryngeal type?

Acknowledgements

- Gahee Yun, Sae Jin Byeon, Yoonjung Kang
- Funded by SSHRC SRG #487010