

Like a Square Peg in a Round Hole: Why Contour Shape Matters for Learning New Intonation Patterns

Mariapaola D'Imperio^{1,2} & James Sneed German¹

¹ Aix-Marseille Université, CNRS, LPL, Aix-en-Provence, France

² Institut Universitaire de France, Paris, France

Introduction

- Imitation across **typologically distinct dialects**
 - Requires approximating phonetic details of phonological structures not in the D1 system
 - D1-D2 imitation is not limited by strong category assimilation as in D1-D1 [1] or high cognitive demand as in L1-L2 [2]
 - Reveals the type and granularity of phonetic detail that can be accessed/implemented by the production system from recently experienced speech events
- D'Imperio & German (2015) [3]:
 - Singapore English (SgE) speakers imitated early F0 peak timing of stress-initial American English (AmE) target words utterance-initially in declaratives
 - No phonetically similar pattern in SgE inventory
- Contour shape marks intonational contrasts in some varieties (Neapolitan Italian [4], German [5],)
- Y/N questions in Singapore English vs. Am. English
 - **Phonetic similarity**: Final rising pattern
 - **Contour shape**: SgE has later rise with no dip (lacks L*), no inflection (i.e, concave versus s-shape)

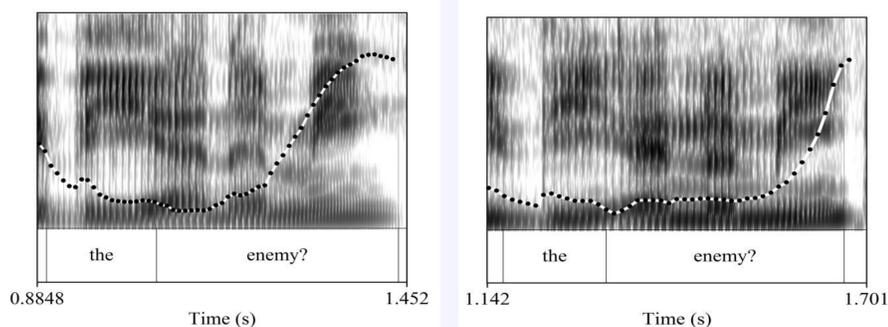


Figure 1: F0 traces for the utterance-final portion of the sentence “*Did the soldiers see the enemy?*” as produced by an American English speaker (left) and a Singapore English speaker (right).

Issues

- **Contour shape**: Can the contour shape of recently experienced tokens be accessed/implemented by the production system?
- **Perceptual assimilation**: Compared to peak timing, does phonetic similarity of Y/N contours limit speakers’ ability to match phonetic detail of AmE targets?

Methods

- Tasks: Baseline reading (D1) + Imitation of AmE speaker
- Target words: trisyllabic, initial stress, sentence-final
- Participants: 18 males, bilingual in SgE/Mandarin
- Measures: Mean F0 of 12 intervals within each target word
- Comparison 1: F0 at timestep X condition
- Comparison 2: Curvature (deg. of inflection) - 3rd coefficient of a 3rd order polynomial regression-fitted to F0 timeseries

References

[1] Cole, J., & Shattuck-Hufnagel, S. (2011). The phonology and phonetics of perceived prosody: What do listeners imitate? *Interspeech 2011* (pp. 969-972). [2] Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of phonetics*, 32(4), 543-563. [3] D'Imperio, M. & German, J. S. (2015). Phonetic detail and the role of exposure in dialect imitation. *Proceedings of ICPHS XVIII*. [4] D'Imperio, M. & Cangemi, F. (2009). The interplay between tonal alignment and rise shape in the perception of two Neapolitan rising accents. *Phonetics and Phonology in Iberia (PaPI) 2009*. [5] Niebuhr, O., D'Imperio, M., Gili-Fivela, B. and Cangemi, F. (2011). Are there “shapers” and “aligners”? Individual differences in signaling pitch accent category. *Proceedings of ICPHS XVII*.
Acknowledgements: This research has been made possible through the support of the A*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the Investissements d'Avenir French Government program, managed by the French National Research Agency (ANR). It was also supported in part by the Erasmus Mundus Action 2 program MULTI of the European Union, grant agreement number 2010-5094-7, and by a grant from the Singapore Ministry of Education Academic Research Fund Tier 1 (2013-T1-002-169).

Results

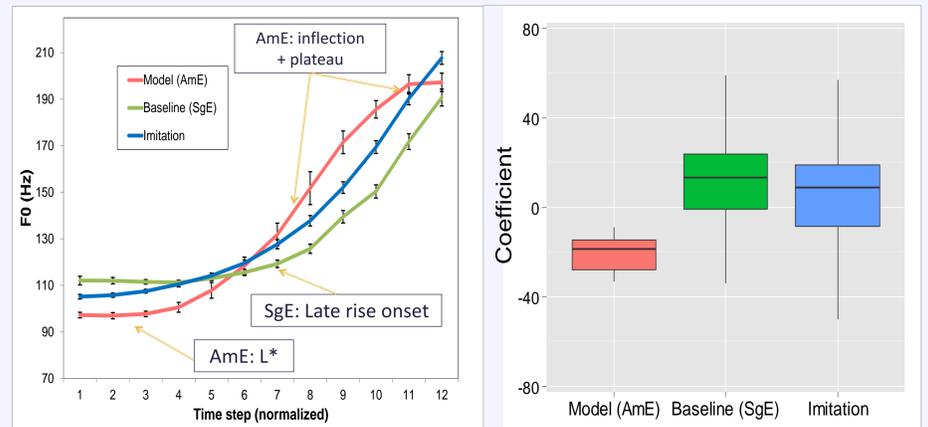


Figure 2: Average F0 by time-normalized increment for AmE model, SgE baseline, and imitation task

Figure 3: Coefficient of third term of regression-fitted polynomial by task (model, baseline, imitation)

- **F0 scaling** (Comparison 1): At each timestep, speakers modified F0 towards the AmE targets (*task x time interaction*: est.max=26.1, tmax=7.83, p<0.0001)
- **Curvature** (Comparison 2): Speakers produced more highly inflected contours during imitation (est.=−4.52, t=2.35, p<0.05), high degree of overlap with baseline
- **Item-by-item accuracy** of curvature depends on self-reported exposure to AmE ($r^2=0.0755$, $p=0.14$)

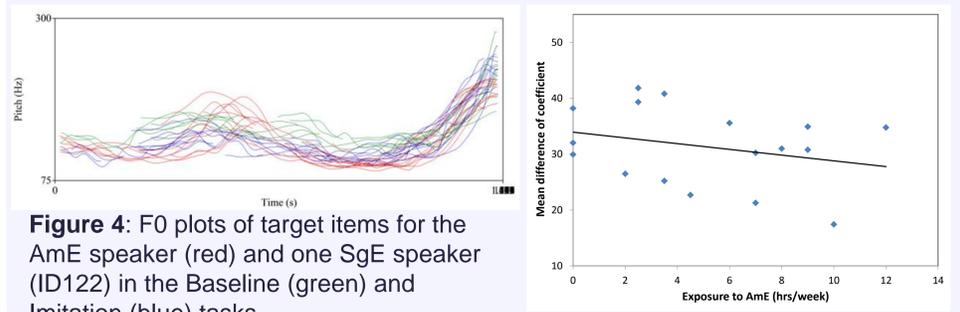


Figure 4: F0 plots of target items for the AmE speaker (red) and one SgE speaker (ID122) in the Baseline (green) and Imitation (blue) tasks

Figure 5: Average by-item difference in coefficient (Imitation-Model) by self-reported weekly exposure to AmE.

Discussion

- **Perceptual assimilation**: Speakers imitated phonetic details of F0 scaling in different regions despite superficial phonetic similarity between D1/D2 contours
- **Contour shape** is not immediately accessible by the production system
 - Speakers may not perceive differences in shape due to effects of perceptual assimilation
 - Producing complex contours may require articulatory practice
- Suggests that imitated tokens are variants of a D1 category whose scaling parameters have been adjusted to provide a better phonetic match to the AmE targets
- Recruitment of native categories does not preclude the imitation of phonetic details as suggested by [1]