

Ethnicity effects in the production of lexical prominence in Singapore English

Adam J. Chong
University of Oxford

Singapore English (SgE) is a contact variety (Grice et al. 2021) spoken alongside at least three other languages – Mandarin Chinese, Malay and Tamil – by a multiethnic (Chinese, Malay and Indian) population. Previous work on Chinese speakers has shown that content words are typically marked on the right edge by a H tone, with the timing of the rise (L* inflection point) tracking the position of lexical prominence. While there is suggestive evidence for speaker backgrounds effects (Chong et al. 2021), existing work on intonation production has primarily examined Chinese speakers (Chong & German 2023). Recent perception work (Chong 2025) has shown that SgE listeners differ in their performance on a stress deafness task (Rahmani et al 2015), with SgE listeners generally showing performance similar to listeners from languages argued to lack contrastive lexical prominence (e.g. French). This was primarily driven by Malay and Indian listeners, with Chinese listeners showing better performance on the task with a wider distribution of accuracy scores. Further, while SgE listeners were generally not sensitive to f0-based cues in word recognition, Chinese SgE listeners showed more sensitivity in certain contexts (Sim et al. 2025). This variability both across and within SgE listener ethnic groups raises the question of the extent to which lexical prominence contrasts are represented across these different groups. In this study, we examine the extent to which SgE speakers from different language/ethnic backgrounds vary in their production of lexical prominence distinctions through an examination of f0.

We present a preliminary analysis of 48 speakers (9 Chinese, 19 Malay, 20 Indian) from a larger corpus of 135 SgE speakers (who also participated in the stress deafness task in Chong 2025). Speakers participated in a remote production task designed to examine the prosodic correlates of focus in four conditions (neutral/broad, contrastive, corrective and wh-focus). 32 trisyllabic target words (e.g. *Germany*) differing in stress (initial vs. medial) were placed in utterance-initial and final positions of carrier sentences. Following Vander Klok et al. (2018), on each trial, speakers saw text of a pseudo-dialogue (1) with an auditorily presented context and a scripted response that they were meant to produce.

- (1) A [context-neutral]: What's in the news? (also auditorily presented)
B [response]: VANILLA will decrease in price.

Here, we focus on target words in the neutral condition in utterance-initial position where the pitch range is largest (381 tokens, Fig. 1). We present a preliminary analysis of mean f0 over each syllable nucleus extracted using a custom Praat (Boersma & Weenink 2022) script, with examination of the dynamic f0 timecourse ongoing. Linear mixed effects models were constructed with mean nucleus F0 as the dependent variable, and fixed factors of stress placement, syllable number (ordered) and SgE speaker group as well as their interactions. Results showed that stress-medial words had a lower f0 on syllable 2 than stress-initial words ($\beta = 9.95$, $SE = 4.01$, $p = 0.015$), conforming to previous results (Chong & German 2023) – f0 starts rising later for stress-medial words. Importantly, there was no significant interaction of speaker group with stress pattern or syllable number. All listeners, regardless of language/ethnic background seem to produce the contrast – with a later rise for stress-medial targets. These findings point to a potential mismatch in perception vs. production of lexical prominence in SgE, especially with the Indian and Malay speakers who previously showed a strong 'stress deafness' effect. Ongoing work will examine if differences emerge in the detailed f0 trajectory over target words and correlations with individual 'stress deafness' scores.

Fig 1. Mean f0 over syllable nucleus by speaker group and stress condition

