It is well known that fundamental frequency (F0) is modulated by talker-specific characteristics such as sex and individual physiology and that F0 can also be affected by speech style (e.g., Loveday 1981). However, even when controlling for these factors, F0 has been shown to differ across languages (Eady 1982, Loveday 1981, Burkhardt 2006, Chan 1998). The aim of the current study is to examine how F0 is affected by the interaction of talker-specific and language-specific characteristics across speech styles in Mandarin, English, and Mandarin-accented English. The addition of a within-talker, between-language condition augments previous work that addresses only comparisons of native speech or native and accented speech; this condition also controls for talker-specific characteristics while varying language. This study is part of a larger research effort aimed at disentangling talker- and language-specific contributions to a wide range of global (i.e. non-contrastive) features of the speech signal in both the spectral and temporal domains.

This study examines the F0 of two male talker populations across three language conditions: L1 English (EE) (N=8); and L1 Mandarin Chinese (MM) and L2 Mandarin-accented English (ME) from the same talkers (N=11). These measurements are from recordings of two tasks in Northwestern’s ALLSSTAR corpus (Archive of L1 and L2 Scripted and Spontaneous Transcripts and Recordings, Bradlow et al. 2011): a short read passage (The North Wind and the Sun, NWS; IPA Handbook 1999) and approximately five minutes of spontaneous speech in response to question prompts (QNA).

Using a linear mixed effects regression, we find main effects of language group and task on the mean ($\beta_{\text{lang}}=-11.9$, $p_{\text{mean}}<0.05$; $\beta_{\text{task}}=-19.6$, $p_{\text{mean}}<0.01$) and standard deviation ($\beta_{\text{lang}}=-29.5$, $p_{\text{sd}}<0.01$; $\beta_{\text{task}}=-34.3$, $p_{\text{sd}}<0.01$) of F0 (see Figure 1). The two-way interaction of language condition and task is also significant for mean ($\beta_{\text{lang:task}}=16.73$, $p_{\text{mean}}<0.05$) and standard deviation ($\beta_{\text{lang:task}}=37.8$, $p_{\text{sd}}<0.05$). Within-language condition comparisons show that the NWS-QNA difference is only significant for MM ($\beta=-19.6$, $p_{\text{mean}}<0.001$); that is, mean F0 is a feature of Mandarin that varies robustly across speech style. Neither English nor Mandarin-accented English vary across style, and there is no significant difference in F0 ranges between English and Mandarin-accented English in either style. This indicates that L2 speakers of English (ME) are able to produce patterns of stylistic variation and F0 ranges that are similar to native English talker ranges for both tasks. In conjunction, there is a significant within-talker correlation of F0 ($r(20)=0.55$, $p<0.01$). Thus, the differences in F0 between Mandarin and Mandarin-accented English can be attributed to language-specific differences between Mandarin and English, and not talker-specific differences.

We conclude that the overall differences between Mandarin-accented English and Mandarin are due to language-specific features. Moreover, only Mandarin shows a significant difference between speech styles, which appear to be a language-specific feature. Thus, there are task-specific strategies that differ across languages as well as general differences across languages in F0. This research begins to disentangle the language- and speaker-specific factors present in the acoustic signal, with evidence that F0 is influenced by language-specific features.
References

Ackerman, L., L. Hesterberg & A. Bradlow. (in prep) Acoustic analysis of Mandarin, English, and Mandarin-accented English using LTASS and F0. Northwestern University.


Figure 1: