THE EFFECT OF LEXICAL FREQUENCY ON TONE PRODUCTION

Yuan Zhao, Dan Jurafsky

Department of Linguistics, Stanford University, CA 94305, U.S.A.
{yuanzhao,jurafsky}@stanford.edu

ABSTRACT

Previous research has identified robust effects on segmental production of lexical factors like word frequency, predictability or neighborhood density. One question that remains unanswered is whether such lexical effects hold also at the suprasegmental level. This study investigates whether lexical factors such as usage frequency affect tone production in Cantonese. We recorded Cantonese monosyllabic words of high and low usage frequency, controlling for segmental factors. The results show that lexical factors do influence suprasegmental production. Words of the same tone but of different usage frequency differ significantly in pitch height. Low-frequency words are hyperarticulated and produced with relatively higher pitch. The overall tone space of low-frequency words is more expanded than that of their high-frequency counterparts.

Keywords: word frequency, tone production, tonal distance, tone space, Cantonese, Chinese

1. INTRODUCTION

Understanding the causes of variation in the surface form of words is crucial to building models of spoken word production. Many such causes of variation have received extensive study, such as the effect of the phonetic context or intonational structure on the realization of phones. While much is thus known about the roles of, e.g., coarticulation or metrical structure on variation, less attention has historically been paid to factors at the level of the word. Recent studies, however, have shown that lexical factors such as word frequency, word predictability, or neighborhood density play an important role in the way a word is realized. Fidelholtz [3], for example, showed that high-frequency (HF) words like *forget* were more likely to have a schwa vowel in the first syllable than low-frequency (LF) words like *forfend*. Recent studies have also found similar effects in natural spoken corpora. Bybee [2] showed that word-final /t/ and /d/ deletion rates in a corpus of spoken Chicano English were higher in HF words than in LF words. Jurafsky [4] confirmed this higher rate of final /t/ and /d/ deletion for HF words in a corpus of American English telephone conversation. They also found that LF words were longer than HF words. Finally, Munson [6] found that LF words had an expanded vowel space compared to HF words (the distance in F1/F2 space of the vowel from a speaker's centroid).

The fact that HF words are more reduced or lenited than LF words has been used to argue for the H&H model [5] or other probabilistic reduction models [4]. These models predict the reduction of HF words at any level of phonetic realization, including the suprasegmental, but this hypothetical impact of frequency on tones has never been tested.

In this study we therefore offer a preliminary investigation of whether lexical factors influence tone production, beginning with this question:

(1) Do speakers use different f0 for LF and HF words (of the same lexical tone)?

In addition, it is possible that there is a tonal analog to the speaker's vowel space, a kind of *tone space* that indicates how distinct each lexical tone is from each other. Frequency may affect this tone space just as it affects vowel space, motivating our second question:

(2) Do speakers make use of an expanded tone space when producing LF words?

2. METHOD

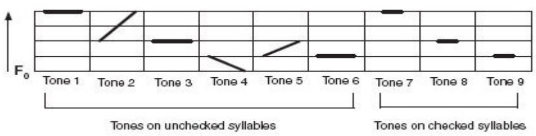
2.1. Overview of methodology

We tested these hypotheses by recording a set of isolated HF and LF monosyllabic words in Cantonese, a 6-toned language. We examined the f0 of each word, testing whether speakers used different f0 values for LF and HF words. We then defined the *tone space* of a speaker, as the average distance of each word to the speaker's tonal centroid. We asked whether LF words were produced with an expanded tone space.

2.2. Test language

Cantonese has six contrastive tones, three level tones that contrast in pitch height (high-level, mid-level and low-level), two rising tones (high-rise and low-rise), that differ in the final pitch target, and a low-falling tone. There are three additional checked tones, i.e. syllables ended with unreleased stops, which mainly differ with the unchecked tones in duration, and hence were not included in the study. Figure 1 shows a schematic representation of the Cantonese tone inventory:

Figure 1: Schematic representations of Cantonese tones [7]



2.3. Materials

The test materials include HF and LF word pairs balanced across different tones. Word token frequency was calculated from the Academia Sinica's Balanced Corpus of Modern Chinese. The mean log frequency of HF words is 3.655 and that of LF words is 0.969. All word pairs are monosyllabic and in either CV or CVC structure. The voicing and manner of the initial consonant is consistent within each pair of high-low frequency words. The vowel length (Cantonese has phonemic long and short vowels) is also consistent within each pair. Only words with nasal codas are used in CVC word pairs and in each pair the same nasal coda is selected. After a set of potential stimuli was chosen, it was given to two native Cantonese speakers to rate the familiarity of the words on a scale of ten. The final stimuli comprise 90 words that are rated above 6 on the familiarity scale.

2.4. Subjects

Eight native Hong Kong Cantonese speakers, four males and four females, participated in the experiment. Subjects' age ranged from 20-52. All were born and raised in Hong Kong and have lived in the United States for less than 3 years.

2.5. Recording and Measurement

The experiments took place in the sound booth of the Phonetics Lab at Stanford University. Subjects were presented with words in a random order one at a time on a computer screen. They produced the words in isolation self-paced without intervention. The sound was recorded with a Panasonic Professional DAT recorder at a sampling frequency of 44.1K and transferred to a computer using the software Audacity. The data were later analyzed with PRAAT [1].

In order to capture as much tonal information as possible, the f0 values of the tonal trajectory of each word were measured at ten equidistant points.

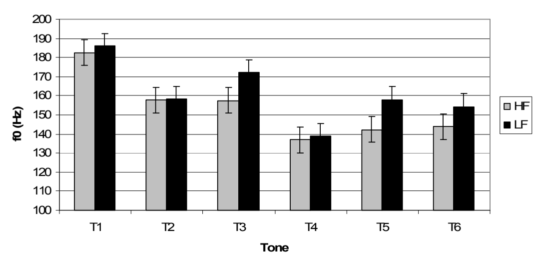
3. RESULTS

3.1. Frequency effect on f0

3.1.1. Mean f0

The mean f0 of six tones of HF and LF words are shown in Figure 2. A two-factor repeated measures ANOVA was used to check whether or not the tone types (tone1, 2, 3, 4, 5, 6) and the frequency category (high vs. low) significantly influence a tone's mean f0. It was found that frequency category does not have a significant main effect on f0, but the interaction between tone types and frequency category has a significant effect on f0 [F (5, 30) =4.965, p<.002]. Post-hoc tests suggested that lexical frequency has a significant main effect on tone3 [F (1, 7) =33.323, p<.001] and tone5 [F (1, 7) =10.677, p<.014]. More specifically, the mean f0 of LF tone3 and tone5 is significantly higher than their HF counterparts (tone3: MD=11.824; tone5: MD=6.709). This result suggests that lexical frequency affects pitch height in tone production, especially on tones in the mid-range (tone3 and tone5). LF words are significantly higher in pitch than HF counterparts.

Figure 2: Mean f0 of HF and LF words of 6 tones



3.1.2. Frequency effect in temporal domain

To further investigate whether the frequency effect is constant along the time course of the pitch trajectory, that is, whether it influences the early part, mid part or the late part of the tonal trajectory, a series of two-factor repeated measures ANOVA was used to check ten f0 points along the trajectory individually. The results show that the interaction of lexical frequency and tone type has a significant effect on the mid part and late part of the pitch trajectory (e.g. 60% time: $F(5, 30) = 3.653$, $p<.011$; 80% time: $F(5, 30) = 4.265$, $p<.005$). In other words, word frequency starts to exert influence on the pitch height of the tone trajectories only about halfway through the word.

4. FREQUENCY EFFECT ON TONE SPACE

Studies on lexical effects on vowel space have shown that vowels of LF words are more dispersed than that of their HF counterparts [6]. If lexical factors are at play on the suprasegmental level as well, we might expect that the overall tone space of LF words is more expanded than that of HF words.

Vowel-space dispersion is usually measured by calculating the mean Euclidean distance from the center of the speakers' F1/F2 space [6]. We propose to measure tone space dispersion parallel to vowel-space dispersion, by calculating the mean Euclidean distance from the center of speakers' f0 space. We thus define the **central f0 at each time point k, $Cf0_k$** as the mean f0 a speaker uses for the time point *k* along the pitch trajectory, averaged across words. For each of the ten equidistant points along the trajectory for every word, we compute $Cf0_k$, the central f0 at the time point *k*, as the average, over all *j* words spoken by the speaker, of $f0^i_k$, the f0 value of word *i* at *k*:

(1) $$Cf0_k = \frac{1}{j}\sum_{i=1}^{j} f0^i_k$$

Given this mean value, we compute the tonal distance in semitones (st) between an f0 value of a particular word *i* at time point *k* and the Central f0 at *k* as follows:

(2) $$\text{Tonal distance to Central f0 (TDC } f0^i_k) = 12\left|\log_2 \frac{f0^i_k}{Cf0_k}\right|$$

In this study, we measured f0 at ten time points over a number of items. Therefore to obtain the averaged tonal distance between a tonal trajectory to Central f0, we need to get the mean of *TDCf0* across all ten time points over all items, as shown by the following:

(3) $$\text{Tonal Dispersion} = \frac{1}{10j}\sum_{k=1}^{10}\sum_{i=1}^{j} TDCf0^i_k$$

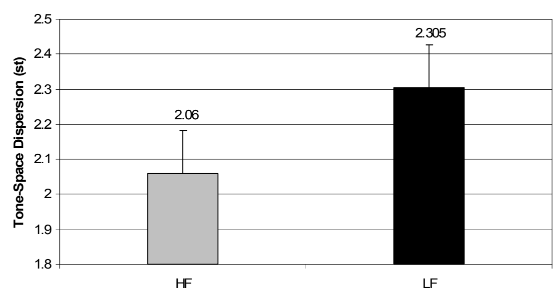
Our hypothesis that lexical frequency affects tonal production predicts that LF words will have a greater tonal dispersion (greater mean distance to the center) than HF words.

4.1. Results

4.1.1 The tone space of Cantonese

In order to exam quantitatively whether or not Cantonese speakers make use of an expanded tone space when producing LF words, we calculated the mean tonal distance to the center of the speaker's f0 space in both HF and LF words. Tone-space dispersion of HF words and LF words is shown in Figure 3.

Figure 3: Tone-Space dispersion of HF and LF words



As suggested by Figure 3, the tone-space of HF words is on average 2.060 semitones dispersed from the acoustic f0 center, while the space of LF words is 2.305 semitones dispersed from the f0 space center. LF words are thus more dispersed than HF words, confirming that speakers tend to use a more expanded f0 space when producing LF words than HF words. A two-factor repeated measures ANOVA with frequency category (high and low) as a within-subject factor and gender as a between-subject factor shows that the usage frequency of a word has a significant main effect on tone-space dispersion ($F(1,6) = 6.910$, $p<.039$). The degree of tone-space dispersion of LF words is on average 0.25 semitone higher than that of HF words. Gender also has a significant effect on the degree of tone-space dispersion ($F(1,6) = 6.640$, $p<.042$, MD=.635 st). Female speakers' tone space expands on average more than that of male speakers, which might be due to the wider pitch range in females.

4.1.2. Three potential competitors

In the Cantonese tone inventory, tone 3, 4 and 6 lie within a similar pitch range and share a slightly falling contour. They are highly confusable in terms of acoustic difference and perceptual distance and therefore are potential competitors in acoustic tone space.

Figure 4: Average F0 of three mid-low tones across high-frequency (left) and low-frequency words (right)

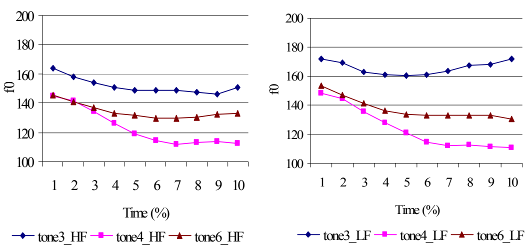


Figure 4 shows that in LF words (right panel), the three tones are more widely distributed in tone space, i.e. are further apart from each other than are the f0 trajectories of HF words (left panel). We measured the acoustic difference among tone competitors by calculating the pair-wise tonal distance between two individual competing tones. The tonal distance between two tones is the summed Euclidean distance between their f0s at all ten points. This gives us an idea of how far apart two tones are distributed.

(4) $$\text{Tonal Distance} = \sum_{i=1...10} \sum_{j=1...10} \sqrt{\left(f0_i - f0_j\right)^2}$$

Figure 5: Tonal distances among mid-low tones across HF and LF words

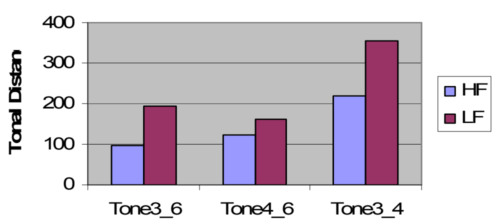


Figure 5 shows greater distance for LF words in every tone pair, indicating that the distribution of those tones is more dispersed in one's tone space for LF words than that of the HF counterparts. The change in the distribution of tones in their tone space when a speaker produces low vs. high frequency words suggests that speakers can dynamically adjust their tone space to increase the distinctness of tones of similar shapes.

5. CONCLUSION

This study shows that lexical factors affect production at the suprasegmental level as well as the segmental level. Words of the same tone but of different usage frequency differ significantly in pitch height. LF words are hyperarticulated and produced with relatively higher pitch. The tone space of LF words is more expanded than that of their HF counterparts; in other words, tones are more dispersed in this acoustic space.

Previous studies on tonal variation mainly focus on how tonal contexts influence variation in the target tone[8], on how neighboring tones trigger tonal assimilation or dissimilation across tonal boundaries. In this study the tone elicitation is done in isolation eliminating possibility of a contextual cause of variation. Our results thus show that at least part of the significant remaining tonal variation in each individual citation tone production can be explained by lexical factors such as lexical frequency.

6. ACKNOWLEDGEMENTS

This work was partially supported by the Edinburgh-Stanford LINK. Thanks to Rebecca Scarborough and Alan Yu for helpful discussions.

7. REFERENCES

[1] Boersma, P. & D. Weenink (2000). *PRAAT: (Version 4.0.30).* http://www.praat.org

[2] Bybee, J. (1999). Usage-based phonology. In M. Darnell et.al (Eds) *Funcionalism and Formalism in Linguistics VI.*

[3] Fidelholtz, J. L. (1975). Word frequency and vowel reduction in English. *CLS* 11, 200-213.

[4] Jurafsky, D., Bell, A., Gregory, M., Raymond, W. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J. & Hopper, P. (ed). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.

[5] Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modeling*, 403-439. Netherlands: Kluwer Academic Publishers.

[6] Munson, B. (to appear). Lexical access, lexical representation and vowel production.

[7] Peng, G. & Wang, W. S. Y. (2004). An innovative prosodic modeling method for Chinese Speech Recognition. *International Journal of Speech Technology* 7, 129-140.

[8] Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics* 25, 61-83.