A FUNDAMENTAL BIAS FOR RESIDUE PITCH PERCEPTION IN TONE LANGUAGE SPEAKERS

Elizabeth M. Petitti, Tyler K. Perrachione

Department of Speech, Language & Hearing Sciences, Boston University, USA
tkp@bu.edu

ABSTRACT

A complex tone composed of only higher-order harmonics elicits a pitch percept equivalent to the tone's missing fundamental frequency (f_0). When judging the direction of residue pitch change between two such tones, however, listeners may have completely opposite perceptual experiences depending on whether they are biased to perceive changes based on the overall spectrum or the missing f_0 (harmonic spacing). Individual differences in residue pitch change judgments are reliable and have been associated with differences in functional neuroanatomy and musical experience. Tone languages put greater pitch processing demands on their speakers than non- tone languages, and we investigated whether these lifelong differences in linguistic pitch processing affect listeners' bias for residue pitch. We found that tone language speakers are significantly more likely to perceive pitch changes based on the missing f_0 than English speakers. These results suggest that tone-language speakers' privileged experience with linguistic pitch fundamentally tunes their basic auditory processing.

Keywords: pitch perception, tone languages, residue pitch, missing fundamental, harmonics

1. INTRODUCTION

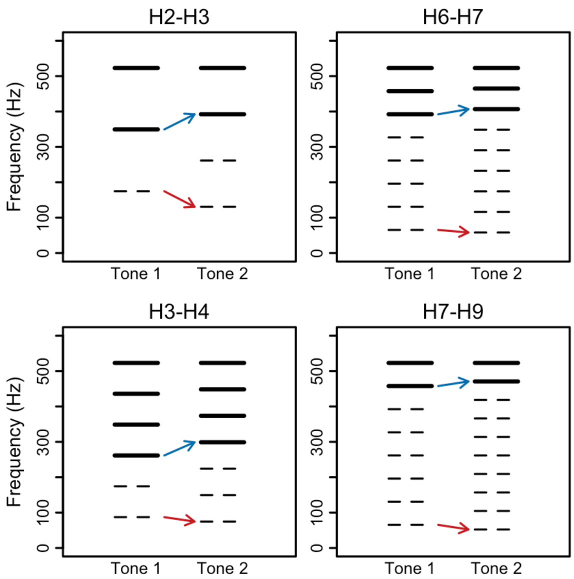
Pitch is used to signal linguistic distinction and pragmatic information, including stress, emotion, and interrogatory intent. Perception of pitch relies on myriad factors, including the fundamental frequency (f_0) and its harmonics. There are individual differences in pitch perception abilities [5], which are sometimes related to differences in lifelong auditory experience, such as musicianship [6]. Linguistic experience also affects the way pitch is perceived, with tone language speakers displaying language-specific perceptual strategies for both speech [11] and non-speech tasks [2]. Tone-language speakers are also better able to perceive pitch under degraded conditions, particularly when stimulus tones are missing key harmonic information [11].

Another impressive demonstration of individual differences in pitch perception stems from judgments about pitch change in pairs of complex tones composed of only higher-order harmonics and missing energy at their f_0. Usually, these tones elicit a pitch percept equivalent to the missing f_0 in a phenomenon known as *residue pitch*. However, when asked to judge the direction of pitch change between two such tones, listeners can give individually consistent yet diametrically opposite answers [10]. The tones illustrated in **Fig. 1** depict this phenomenon. Listeners who perceive residue pitch change based on the spacing between harmonics (*i.e.,* the missing f_0) will respond that the change in the tones' pitch moves downward. Conversely, listeners who perceive the residue pitch change based on the global shift in frequencies of the present partials will respond that the change in pitch moves upwards. Certain acoustic factors reliably affect listeners' residue pitch change judgments regardless of their individual fundamental vs. spectral bias. For example, increasing implicit f_0 difference, number of harmonics, harmonic resolution, and stimulation amplitude have all been shown to generate more f_0-biased responses [3,5,7].

Despite these consistent acoustic effects, individual differences persist. Neuroanatomic differences may relate to residue pitch bias, with f_0-listeners displaying increased gray matter in left Heschl's gyrus and greater left-lateralized auditory evoked potentials [5]. Some studies have found greater fundamental bias in musicians [7], whereas others have not found this effect [3,5].

In addition to these differences in pitch perception, long-term linguistic experience with pitch has also been shown to affect pitch perception abilities [2,11]. Given the increased demands of pitch perception for speakers of tone languages, we hypothesized that tone language speakers will be more sensitive to the consistent harmonic relationships associated with a specific f_0 [8] and will therefore have a greater fundamental bias when judging residue pitch change in pairs of missing-f_0 tones. Here, we asked native speakers of English and tone languages to make residue pitch change judgments for pairs of missing-f_0 complex tones. Our results indicate that tonal language listeners privilege f_0 information significantly more than English listeners when attending to this basic, non-linguistic, auditory processing task.

Figure 1: Stimulus parameterization for residue pitch change judgment task. Four example stimulus pairs are shown, titled with the shift in lowest present harmonic. In the top left panel, Tone 1 is composed of two present harmonics (solid lines) at 349 and 523Hz with an implicit f_0 of 174Hz; Tone 2 is also composed of two harmonics, now at 392 and 523Hz, and with an implicit f_0 of 131Hz. Red arrow: residue pitch change percepts based on missing f_0 (fundamental listeners). Blue arrow: residue pitch change percepts based on the upward shift in frequencies of the present partials (spectral listeners).



2. METHODS

2.1. Subjects

Two groups of participants completed this experiment: native tone language speakers and native English speakers. The tone language group (N=40, 10 male and 30 female, age 18-28 M=21.4 years) was comprised of native speakers of Mandarin (N=21), Cantonese (N=5), bilingual Mandarin/Cantonese (N=11), and Vietnamese (N=3). The English language group (N=40, 9 male and 31 female, age 18-26, M=20.4years) had no prior exposure to a tone language. All participants demonstrated normal hearing by passing a basic audiometric screening and had a self-reported history free from speech, language, or hearing difficulties. All included participants exhibited accurate pitch perception judgments [9] with >90% performance on a control task (see below); 31 additional participants were recruited but not included because their pitch judgments were not reliable (<90% control task accuracy).

2.2. Stimuli

We synthesized 72 pairs of harmonic complex tones with missing-f_0 components. Following Schneider *et al.* [5], we assessed residue pitch change judgments across a range of acoustic factors by parametrically varying the number of present harmonics, the frequency of the highest present harmonic, and the harmonic order. Examples of these manipulations are depicted in **Fig. 1**. The number of present harmonics (2, 3, or 4) and the frequency of the highest-order harmonic (293, 523, 932, 1661, 2960, or 5274Hz) were kept constant across the two tones in a pair. This allowed the difference in implicit f_0 to be determined by the harmonic order [3]. The harmonic order between the two tones differed such that the lowest present harmonic changed (from H2-H3, H3-H4, H6-H7, or H7-H9) while the frequency of the highest was unchanged. These manipulations resulted in tones with an implicit f_0 of 24-1758Hz, present frequency components of 146-5274Hz, and mean spectra of 212-4977Hz.

During stimulus generation, a rounding error occurred causing 36 complex tones to be synthesized without the intended highest harmonic, affecting the harmonic relationship between 36 tone pairs. As a result, 72 trials were excluded from the analysis, and the results below are based on *only* the 72 trials where tone pairs had the correct harmonic composition.

In order to insure that participants were making accurate, authentic judgments about their perceptual experiences of pitch change [9], we included control tones with unambiguous pitch differences that matched the spectral composition and implicit f_0 of experimental tones. Control stimuli consisted of 12 pure tones (f_0=195-4102Hz) and 12 complex tones in which both f_0 and all higher-order harmonics were present (f_0=37-1055Hz, H1-H12 present).

All stimuli were synthesized in Praat with a sampling rate of 44.1kHz at 16bits. Each tone was presented at 50dB SPL for 500ms, including 10ms linear rise-fall times, and a 250ms silent interval between tones in a pair.

2.3. Procedure

Participants were seated in a sound-attenuated chamber. Stimuli were delivered over Sennheiser HD380 Pro circumaural headphones via a Behringer FCA1616 USB audio interface, controlled by PsychoPy (v1.80.0). Participants first completed a basic audiometric screening in each ear consisting of octave-spaced pure tones from 1000-4000Hz at 20dB HL. Participants learned the demands of the experimental task by completing 24 practice trials consisting of complex and pure tone pairs with unambiguous pitch changes. Participants received automatic feedback on the accuracy of their pitch judgements. None of the practice tones were included in the experiment.

The pitch-change judgment task was based on previous designs examining missing-f_0 tones [3,5,7]. The experimental trials were broken into two runs of 72 trials each, separated by a self-paced break. Tone pairs were presented in both rising-f_0 and falling-f_0 orders, counterbalanced across the two runs, so that there was no design bias in the direction of implicit f_0 change. Participants responded via keyboard, pressing the "up" arrow for perceived rising pitch and the "down" arrow for perceived falling pitch. The experiment was self-paced, and lasted approximately 30 minutes.

2.4. Data analysis

Listeners' response to each trial was assigned a score of "0" to indicate a spectral pitch judgment (*i.e.*, rising harmonic frequencies → rising pitch percept) or a score of "1" to indicate a fundamental pitch judgment (*i.e.*, more closely spaced harmonics → falling pitch percept). Overall listener bias (*i.e.*, the probability of a fundamental pitch judgment, P(f_0)) was calculated by computing the average of pitch judgment scores, yielding a number between 0 (completely spectral bias) and 1 (completely fundamental bias). Trials with response times exceeding two standard deviations from a participant's mean were excluded from analysis.

Inferential statistics on participant responses were conducted using generalized linear mixed effects models for binomial data with fixed factors including Group (English vs. Tone language) and number of present harmonics, and a maximal random effects structure [1] including within-participant intercepts and slopes and within-stimulus intercepts.

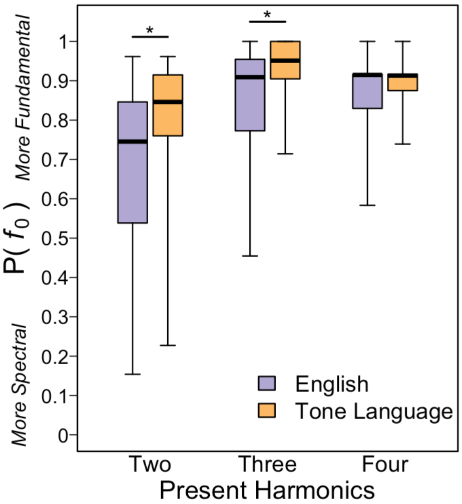
3. RESULTS

3.1. Linguistic factors affecting P(f_0)

Aggregating over all experimental trials, English listeners' P(f_0) raged from 0.39-0.96 (M=0.80). Tone-language listeners' P(f_0) ranged from 0.61-0.96 (M=0.88). Unlike previous studies [3,5], no participants exhibited completely spectral (P(f_0)=0) or completely fundamental (P(f_0)=1) perceptual biases. Tone-language speakers displayed significantly greater P(f_0) perceptual bias than native English speakers (z=3.25, p=0.0016). This difference was retained across stimuli containing two (z=3.35, p=0.0004; Cohen's d=0.78) and three (z=3.16, p=0.0016; d=0.66), but not four (z=0.99, p=0.32; d=0.21) present harmonics (**Fig. 2**).

English listeners' P(f_0) for two-harmonic stimuli ranged from 0.15-0.96 (M=0.68), from 0.45-1.0 (M=0.87) for three, and from 0.58-1.0 (M=0.87) for four. Tone-language listeners' P(f_0) ranged from 0.23-0.96 (M=0.81) for two, from 0.71-1.0 (M=0.93) for three, and from 0.74-1.0 (M=0.89) for four present harmonics.

Figure 2: P(f_0) is significantly higher for tone speakers with two and three present harmonics.



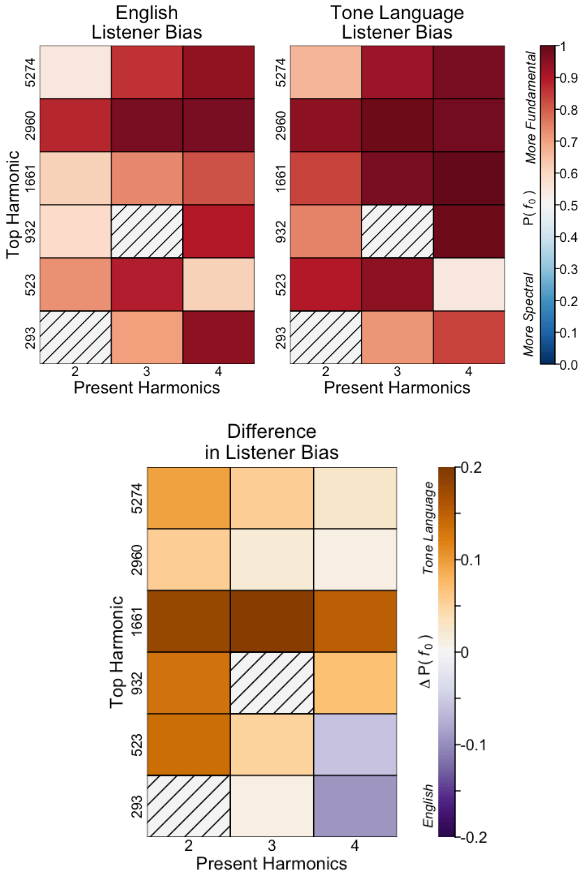
3.2. Acoustic factors affecting P(f_0)

Consistent with previous studies [3,5,7], both listener groups demonstrated increasing P(f_0) as the number of harmonics increased: three-harmonic stimuli were significantly more fundamental than two-harmonic (z=2.72, p=0.0065); however, P(f_0) did not differ between three- and four-harmonic stimuli (z=0.24, p=0.81). The trend for increasing P(f_0) with increasing acoustic information was consistent across both groups (no language × harmonic composition interaction; z=0.41, p=0.68). These relationships are depicted in **Fig. 3**.

2.6. Musical factors affecting P(f_0)

Given research demonstrating effects of musical training on pitch perception [4], including residue pitch perception differences between musicians and nonmusicians [7], we examined the impact of musical training on P(f_0) in the native English participants. We found no evidence that number of years played (r=-0.012, p=0.94) nor start age (r=0.058, p=0.71) impacted P(f_0), corroborating other studies that found no effect of musical training on pitch perception bias [3,5].

Figure 3: P(f_0) for tone-language and English speakers as a function of number of present harmonics and frequency space (top). Greater P(f_0) values (dark red) are found for increasing number of harmonics. Differences between language groups (ΔP(f_0), bottom) reveal systematically greater f_0 residue pitch perception in tone language speakers relative to English speakers. There are two null cells due to no included stimuli with these features.



3. DISCUSSION

We investigated differences in residue pitch perception between native speakers of English and tone languages. Native speakers of tone languages exhibited a consistent and significantly greater bias to perceive residue pitch changes between pairs of complex tones based on the change in implicit f_0 (harmonic spacing) than native English speakers. This language-based difference was consistent across a variety of acoustic factors that independently affect P(f_0). A greater bias to perceive residue pitch changes based on the implicit f_0 may be the result of lifelong increased pitch processing demands imposed on speakers of a tone language, for whom subtle pitch differences signal semantic distinctions between words. Moreover, we found no evidence that listener bias for tones with residue pitch extends to musical expertise, suggesting linguistic experience impacts basic auditory processing in a way that other individual factors do not. The privileged processing of pitch in tone languages may facilitate the association between the co-occurrence of harmonic partials and f_0 [8], such that even when f_0 is absent, the harmonic partials alone are sufficient to evoke an implicit f_0-based percept of pitch.

These data suggest that lifelong experience with processing linguistic pitch in a tone language results in an auditory system that is fundamentally tuned to different features of even non-linguistic, basic acoustic stimuli compared to speakers of a non-tone language. These results further emphasize the importance of considering the linguistic demands of audition when developing models of basic auditory processing.

7. REFERENCES

[1] Barr, D., Levy, R., Scheepers, C., Tily, H. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang*. 68, 255–278.

[2] Bent, T., Bradlow, A., Wright, B. 2006. The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *J. Exp. Psychol. Human.* 32, 97–103.

[3] Ladd, D. *et al.* 2013. Patterns of individual differences in the perception of missing fundamental tones. *J. Exp. Psychol. Human.* 39, 1386–1397.

[4] Perrachione, T. K. *et al*. 2013. Evidence for shared cognitive processing of pitch in music and language. *PloS One*. 8, e73372.

[5] Schneider, P. *et al*. 2005. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* 8, 1241-1247.

[6] Schön, D., Magne, C., Besson, M. 2004. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*. 41, 341–349.

[7] Seither-Preisler, A. *et al.* 2007. Tone sequences with conflicting fundamental pitch and timbre changes are heard differently by musicians and nonmusicians. *J. Exp. Psychol. Human.* 33, 743-751.

[8] Shamma, S., Klein, D. 2000. The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *J. Acoust. Soc. Am*. 107, 2631–2644.

[9] Semal, C., Demany, L. 2006. Individual differences in the sensivity to pitch direction. *J. Acoust. Soc. Am.* 120, 3907–3915.

[10] Smoorenburg, G. 1970. Pitch perception of two-frequency stimuli. *J. Acoust. Soc. Am.* 48, 924-942.

[11] Stagray, *et al.* 1992. Contributions of the fundamental, resolved harmonics, and unresolved harmonics in tone-phoneme identification. *J. Speech. Hear. Res.* 35, 1406-1409.