# Pitch processing is shared between language and music

**Tyler Perrachione[1], Ev Fedorenko[1], Louis Vinke[2], Edward Gibson[1] & Laura Dilley[3]**
[1]Massachusetts Institute of Technology, [2]Massachusetts General Hospital, [3]Michigan State University

## Background and motivation

Language and music share many features, including reliance on **pitch**.

***Pitch in language:*** helps disambiguate syntactic structures (e.g., Beach, 1991; Krajlic & Brennan, 2005), conveys aspects of pragmatic and semantic meaning (e.g., Fromkin, 1978; Breen et al., 2010).
***Pitch in music:*** melodies are encoded through the patterns of discrete pitches played on instruments or sung.

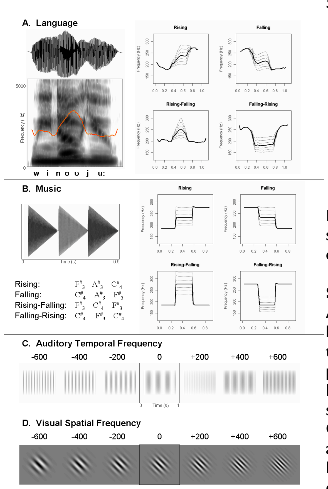
**Previous evidence** suggests that pitch processing may be shared between language and music:
- in the auditory brainstem linguistic pitch patterns are encoded more robustly in musicians than non-musicians (Wong et al., 2007)
- musicians are better at perceiving spoken linguistic input in noise conditions, a process that depends on following the pitch pattern of the attended voice (Parbery-Clark et al., 2009)
- individuals with musical training are better at learning languages that use pitch as a phonological contrast (Wong & Perrachione, 2007)
- musically tone-deaf (amusic) individuals are also impaired in making linguistic distinctions based on pitch (e.g., Patel et al., 2008; Hutchins et al, 2010)

**Limitations of previous evidence**:
- many studies focus on expert musicians
- some of the evidence can be explained without postulating shared pitch-processing machinery (e.g., pitch perception in language and music may both depend on the same low-level sensory pathways - e.g., in the auditory brainstem - or on the same domain-general attention / working memory / motivation mechanisms)
- some evidence for distinct pitch mechanisms: left-lateralized system for linguistic pitch processing for semantic content and right-lateralized system for the processing of musical melody or sentence prosody (e.g., Wong, 2002; Xu et al., 2006; cf. Luo et al., 2006)

**Current study:** evaluates the hypothesis that pitch processing in language and music is shared above and beyond these abilities' mutual reliance on low-level sensory-perceptual pathways or domain-general processes like attention or motivation.

## Methods

### Example psychophysical stimuli




**Design:**
5 conditions:
- pitch processing in linguistic stimuli
- pitch processing in musical stimuli
- control condition #1: a non-linguistic/non-musical test of psychophysical pitch discrimination threshold (controls for basic sensory acuity in pitch discrimination)
- control condition #2: a test of temporal frequency discrimination (controls for basic, non-pitch, auditory perceptual acuity)
- control condition #3: a test of visual spatial frequency discrimination (controls for individual differences in attention and motivation)

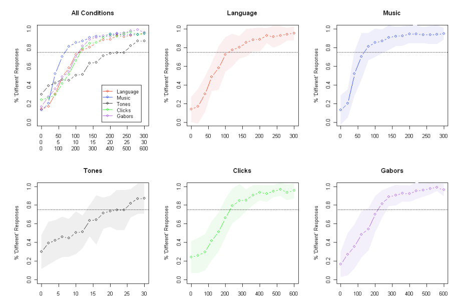
**Key prediction:** If pitch processing in language and music is supported by shared mechanisms, then a relationship between the two should remain after controlling for shared sensory and domain-general factors.

**Stimuli details** (see the manuscript for additional details):
A: Waveform and spectrogram of linguistic stimulus with phonetic alignment; the four linguistic pitch contours (dark trace) showing +/- 100, 200, and 300 cents deviants (light traces). The rising-falling pitch contour is overlaid on the spectrogram (red trace) to show phonetic alignment.
B: Waveform and notation of musical stimulus; the four pitch contours (dark trace) showing +/- 100, 200, and 300 cents deviants (light traces).
C: Spectrograms of auditory temporal frequency stimuli, with template stimulus in box and example +/-200, 400 and 600 cents deviants.
D: Visual spatial frequency stimuli ("Gabor patches"), with template stimulus in box and example +/-200, 400 and 600 cents deviants.

17 participants (none with extensive musical training)

## Results

### Discrimination contours across stimulus conditions




### Task performance by condition

| | Overall Accuracy | | $d'$ | | Threshold (cents) | |
|---|---|---|---|---|---|---|
| | Mean | Std. Dev. | Mean | Std. Dev. | Mean | Std. Dev. |
| Language | 0.77 | ± 0.09 | 2.01 | ± 0.86 | 151 | ± 74 |
| Music | 0.83 | ± 0.09 | 2.33 | ± 0.89 | 129 | ± 92 |
| Tones | 0.65 | ± 0.09 | 1.05 | ± 0.77 | 26 | ± 5 |
| Clicks | 0.75 | ± 0.06 | 1.54 | ± 0.66 | 313 | ± 137 |
| Gabors | 0.79 | ± 0.08 | 2.02 | ± 1.05 | 296 | ± 143 |

### Linear models

| Response | Terms | Overall Accuracy β | Overall Accuracy $p_{term} <$ | $d'$ β | $d'$ $p_{term} <$ | Threshold β | Threshold $p_{term} <$ |
|---|---|---|---|---|---|---|---|
| Language | Music | 0.655 | 0.000 | 0.525 | 0.005 | 0.599 | 0.001 |
| | Tones | 0.181 | 0.191 | 0.024 | 0.949 | 2.867 | 0.220 |
| | Clicks | 0.261 | 0.165 | 0.723 | 0.112 | 0.046 | 0.595 |
| | Gabors | -0.023 | 0.822 | -0.087 | 0.675 | 0.004 | 0.956 |
| | $R^2 =$ | 0.92 | | 0.73 | | 0.85 | |
| | $p_{model} <$ | $5.7\times10^{-7}$ | | 0.0012 | | $3.5\times10^{-5}$ | |
| Music | Language | 1.140 | 0.000 | 0.887 | 0.005 | 1.015 | 0.001 |
| | Tones | -0.090 | 0.635 | 0.227 | 0.639 | -0.628 | 0.841 |
| | Clicks | -0.210 | 0.415 | -0.687 | 0.256 | 0.093 | 0.406 |
| | Gabors | 0.060 | 0.682 | 0.127 | 0.637 | 0.022 | 0.826 |
| | $R^2 =$ | 0.87 | | 0.57 | | 0.83 | |
| | $p_{model} <$ | $9.3\times10^{-6}$ | | 0.02 | | $5.9\times10^{-5}$ | |

### Pairwise correlations

* = significant at Bonferroni-corrected alpha=.00167

| Conditions | | Overall Accuracy $r =$ | Overall Accuracy $p <$ | $d'$ $r =$ | $d'$ $p <$ | Threshold $r =$ | Threshold $p <$ |
|---|---|---|---|---|---|---|---|
| Language | Music | 0.927 | 0.000 * | 0.720 | 0.001 * | 0.905 | 0.000 * |
| | Tones | 0.780 | 0.000 * | 0.660 | 0.003 | 0.601 | 0.008 |
| | Clicks | 0.749 | 0.000 * | 0.693 | 0.001 * | 0.661 | 0.003 |
| | Gabors | 0.389 | 0.111 | 0.550 | 0.018 | 0.562 | 0.015 |
| Music | Tones | 0.671 | 0.002 | 0.433 | 0.073 | 0.522 | 0.026 |
| | Clicks | 0.626 | 0.005 | 0.378 | 0.122 | 0.684 | 0.002 |
| | Gabors | 0.374 | 0.126 | 0.376 | 0.124 | 0.550 | 0.018 |
| Tones | Clicks | 0.752 | 0.000 * | 0.881 | 0.000 * | 0.330 | 0.180 |
| | Gabors | 0.384 | 0.115 | 0.784 | 0.000 * | 0.540 | 0.021 |
| Clicks | Gabors | 0.425 | 0.079 | 0.788 | 0.000 * | 0.543 | 0.020 |

## Conclusions

We observed a significant and strong relationship between individuals' pitch processing abilities in language and music even after accounting for the contribution of basic non-linguistic and non-musical sensory acuity, including auditory acuity for pitch, and domain-general mnemonic, attentional, and motivational factors that bear on laboratory tests of perception.

These data support the hypothesis that **cognitive mechanisms for pitch processing in language and music are shared beyond simple reliance on overlapping auditory sensory pathways or domain-general attention or working memory.**

Importantly, the relationship between pitch processing in language and music holds in the general population, not in trained musicians or individuals with neurological disorders (e.g., amusia).

**There exist higher-level cognitive and neural mechanisms that support the processing of pitch information across domains.**

**Corresponding authors:**
**Ev Fedorenko:** evelina9@mit.edu **Tyler Perrachione:** tkp@mit.edu