

Earlier music biases subsequent musical instrument categorization

Joshua M. Lanning and Christian E. Stilp

Department of Psychological and Brain Sciences, University of Louisville



INTRODUCTION

Perception of a sound is heavily influenced by surrounding sounds. When spectral properties differ between earlier (context) and later (target) sounds, this can produce **Spectral Contrast Effects (SCEs)** that bias categorization of later sounds.

Context	More likely to perceive
Musical segment (unmodified)	"French horn" or "tenor saxophone"
Musical segment with French horn frequencies emphasized	"tenor saxophone"
Musical segment with tenor saxophone frequencies emphasized	"French horn"

Earlier studies have found that SCEs influence perception of speech (Stilp *et al.*, 2015) as well as music (Stilp *et al.*, 2010). The magnitudes of SCEs are not all-or-none, but vary continuously for speech (Stilp *et al.*, 2015; Stilp & Assgari, 2017) and music (Frazier *et al.*, under review). Recently, SCEs in speech perception were produced by unfiltered sentence contexts that already had the desired spectral properties without any additional filtering (Stilp & Assgari, under review).

The present study tested whether unfiltered contexts (*i.e.*, musical segments that already have the desired spectral properties) produce SCEs in perception of music in the same manner as was reported for speech.

STIMULI

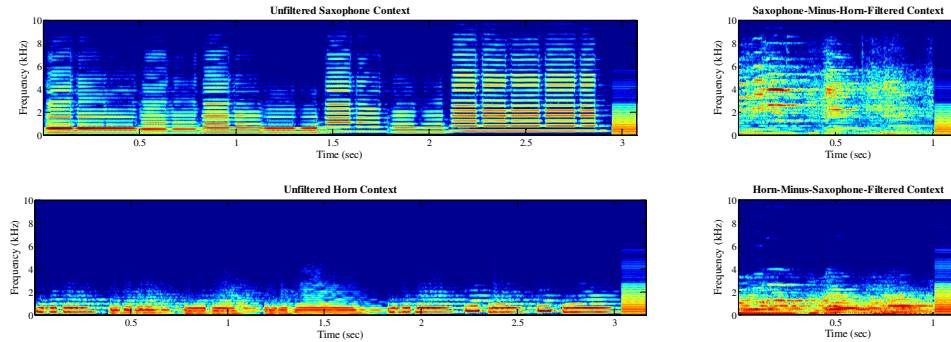
Targets

Target stimuli were drawn from the McGill University Musical Samples database (Opolko & Wapnick, 1989)

- Same stimuli as tested in Stilp *et al.* (2010) and Frazier *et al.* (under review)
- Both instruments were 196 Hz (G3) and 140 ms in duration
- Stimuli were proportionately mixed so that the amplitude of one instrument was (+30, +18, +6, -6, -18, -30) dB relative to the other
- Stimuli with 30-dB differences served as the series endpoints

Contexts

- Unfiltered
 - Tenor Saxophone (<https://tinyurl.com/Lanningsaxophone>)
 - From YouTube video of Ivan Renta playing for *Profiles in Greatness*; 2938 ms
 - French Horn (<https://tinyurl.com/Lanninghorn>)
 - From YouTube video of Sarah Willis of the Berlin Philharmonic; 3032 ms
 - Excerpt of the seldom used unstopped playing style to match the unstopped French horn target stimulus
- Filtered
 - One second excerpt from Franz Schubert's string quintet in C major, Allegretto (taken from compact disc)
 - Musical context processed by spectral envelope difference filters (created by taking the spectra of the two instrument endpoints and subtracting them from each other; Watkins, 1991)
 - Spectra calculated using 512-point Fourier transforms, smoothed with 256-point Hamming windows with 50% overlap
 - Same filters as used in previous studies (Stilp *et al.*, 2010; Frazier *et al.*, under review)



METHODS

Participants

- 22 undergraduate students from the University of Louisville with self-reported normal hearing

Procedure

- Exposure:** Each musical instrument endpoint stimulus (French horn or tenor saxophone) was played twice along with its verbal label
- Practice:** 100 trials with musical endpoint stimuli; $\geq 90\%$ accuracy required to continue to test session
- Main Experiment:** 120 trials (in random order) in each of two blocks (in counterbalanced orders)
 - One block presented trials with unfiltered musical contexts; the other block presented trials with filtered musical contexts
 - Trial structure: musical context immediately followed by target instrument sound, which listeners identified as "French horn" or "tenor saxophone"
 - Performance criterion: must maintain $\geq 80\%$ accuracy on endpoint stimuli across blocks (9/22 failed to meet this criterion; their data excluded from analyses)
- Post-Test Survey** (same survey as administered in Stilp *et al.*, 2010 and Frazier *et al.*, under review)
 - Question 1: Please rate your musical performing ability: 1 (none), 2, 3, 4, 5 (virtuoso)
 - Question 2: How many years of solo musical performance experience (formal training/instruction) have you had?
 - Question 3: How many years of ensemble musical performance experience (formal training/instruction) have you had?
 - Question 4: Do you have any other relevant musical experience to share?
 - Question 5: Did you recognize the musical selection in the final testing block? Can you name it?

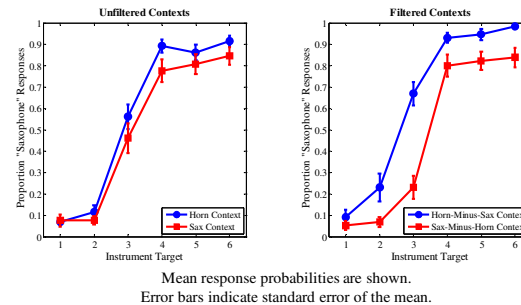
RESULTS

SCE was defined as the change in % "saxophone" responses to target stimuli across conditions (higher %s following horn / horn-minus-saxophone contexts)

- Unfiltered Contexts: mean SCE = 6.15% shift
 - One-tailed *t*-test against 0: $t_{12} = 2.88, p < .01$
- Filtered Contexts: mean SCE = 17.31% shift
 - One-tailed *t*-test against 0: $t_{12} = 7.88, p < .001$
- SCEs were significantly larger following Filtered Contexts
 - Paired-samples *t*-test: $t_{12} = 4.28, p < .01$

Musical experience was not correlated with SCE magnitudes

- Question 1 (Spearman's correlations)
 - Unfiltered SCEs: $\rho = -0.46, p = 0.11$
 - Filtered SCEs: $\rho = -0.18, p = 0.56$
- Question 2
 - Unfiltered SCEs: $r = -0.45, p = 0.13$
 - Filtered SCEs: $r = -0.27, p = 0.37$
- Question 3
 - Unfiltered SCEs: $r = -0.42, p = 0.16$
 - Filtered SCEs: $r = -0.41, p = 0.16$



DISCUSSION

Unfiltered music significantly affected categorization of the subsequent musical instrument target

- Extends the findings of Stilp & Assgari (under review) in speech to perception of musical instrument sounds
- Further extends the parallels between speech SCEs and music SCEs listed in the Introduction

SCEs produced by the Unfiltered context were significantly smaller than SCEs produced by the Filtered context

- Consistent with speech results in Stilp & Assgari (under review)
- Specific reason for this difference is unclear, but Unfiltered block presented two different pieces of music playing different notes on each trial, creating a lot of variability; Filtered block presented the same piece of music on every trial, resulting in substantially less variability

Surprisingly low percentage of participants met the modest performance criterion of $\geq 80\%$ correct on endpoint stimuli (13/22 or 59%)

- Possible explanations include:
 - Confusion resulting from responding to some horns and saxophones (targets) but not others (contexts)
 - Identifying endpoint stimuli by themselves in practice block is easier than identifying them following a musical context

The post-test survey showed no relationship between musical experience and SCEs in musical instrument categorization

- This participant sample did not have much musical experience
 - Question 1: mean = 1.77, median = 1
 - Question 2: mean = 1.08, median = 0
 - Question 3: mean = 0.85, median = 0
- Musical experience was not used in recruitment of participants
- Thus, these results do not demonstrate a lack of influence of musical experience on context effects, just that it did not correlate with responses for this sample
- Future studies should recruit experienced musicians and repeat this experiment

Filtered results replicate Stilp *et al.* (2010) and Frazier *et al.* (under review); Unfiltered results extend these context effects to listening conditions that more closely resemble everyday perception

REFERENCES

- Frazier J, Assgari AA, & Stilp CE (under review) *Attention, Perception, & Psychophysics*
- Opolko F & Wapnick J (1989) *McGill University Master Samples user's manual*.
- Stilp CE, Alexander JM, Kiefe M, & Klueder KR (2010) *Attention, Perception, & Psychophysics*, 72(2), 470-480
- Stilp CE, Anderson PW, & Winn MB (2015) *Journal of the Acoustical Society of America*, 137(6), 3466-3476
- Stilp CE & Assgari AA (2017) *Journal of the Acoustical Society of America*, 141(2), EL153-EL158
- Stilp CE & Assgari AA (under review) *Attention, Perception, & Psychophysics*
- Watkins AJ (1991) *Journal of the Acoustical Society of America*, 90(6), 2942-2955