School of Brain and Behavioral Science

Real-Time Modulation Perception

in Western Classical Music

Brendon Mizener

THE UNIVERSITY OF TEXAS AT DALLAS

Outline

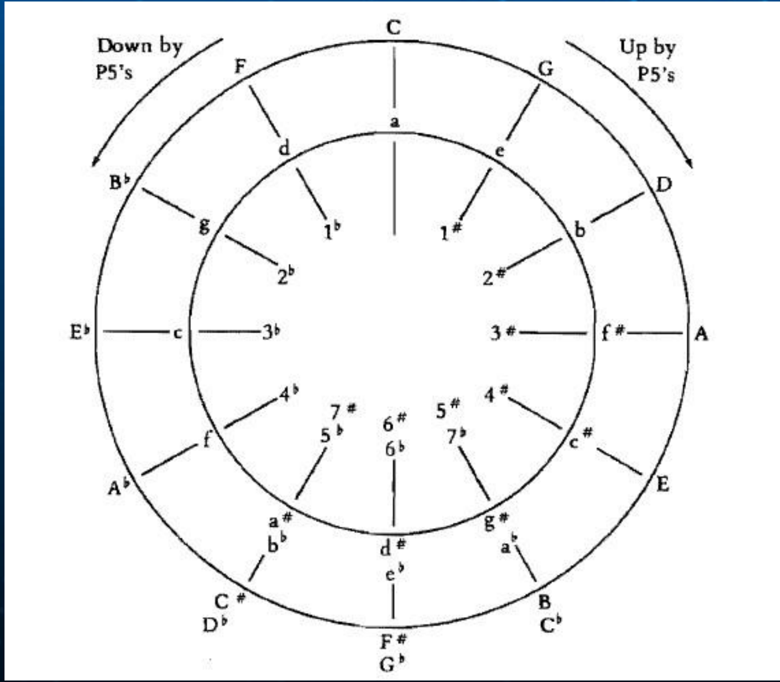
- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Outline

- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Background: Music Theory

- Key words:
  - Tonic, Dominant
  - Key, key area
  - Functional harmony
  - Mode
  - Relative & parallel minor



Benjamin, Horvit, and Nelson (2003)

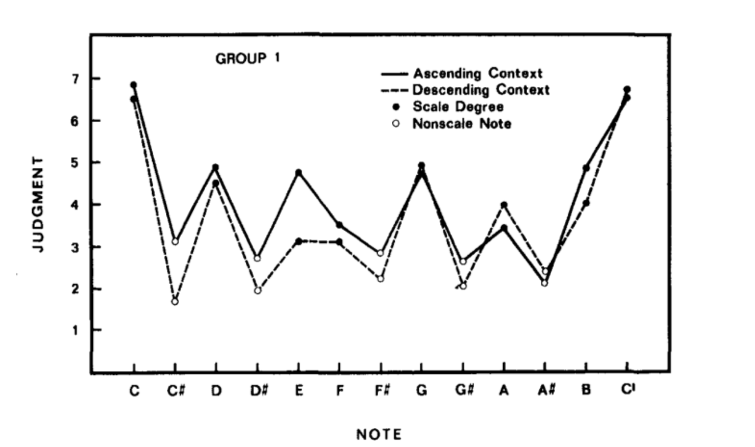
Background: Music Perception

- Music perception is multi-dimensional:
  - Pitch class & octave (Shepard, 1982)
  - Key & key membership (Krumhansl & Shepard, 1979)
  - Intervals & note relationships (Dowling, 1978)
  - Rhythm & Temporal Expectancy (Narmour, 2015)
- Perception and understanding of tonic is central to music perception. (Krumhansl & Kessler, 1982)

Tonal Hierarchy Model (Krumhansl & Shepard 1979)

- Notes in a key are hierarchical:
  - Tonic is most important, followed by 5th and 3rd scale degrees
  - Other notes in the key outside of the tonic triad are less important
  - Out of key notes are least important

- Issues with this model:
  - Assumption of a priori knowledge of tonic (Butler, 1983)
  - Ecological validity of stimuli (Vuvan, Prince, & Schmuckler, 2011)
  - Reference pitch



(Krumhansl & Shepard 1979)

Other Models of Tonality Induction

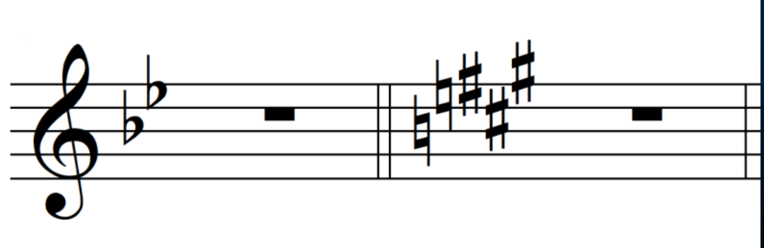
- Rare intervals hypothesis (Butler 1989)

(Butler 1989)

- Tonal Decay model (Huron & Parncutt 1993)

What is a modulation?

- Process by which a composer changes tonic in a composition.
- Many distinctly defined types of modulation in western classical music.
- Adds interest and structure to a piece of music



Perceiving the Tonic in Motion

- Listeners, regardless of training, follow modulations with a high degree of accuracy. (Cuddy & Thompson 1992)
- More recent harmonic material informs key perception to a greater degree than older material. (Krumhansl & Kessler 1982)
- Closely related keys are incorporated into perception more quickly than distantly related keys. (ibid.)
- Our perception of key seems to be dynamic. (Toivianen & Krumhansl 2003)
- Pitch distance and harmonic difference affect discrimination and response bias differently. (Kleinsmith & Neill 2017)
- Both training and enculturation affect accuracy in perception of modulations. (Raman & Dowling 2017)

Outline

- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Questions

1. Do music listeners passively retain information on key region independent of topical, salient features of the music?
2. To what extent does training affect the storage, processing, and access to that information, if it exists?
3. What topical features influence our understanding of key regions and the movement between them?
4. What is the balance between melodic and harmonic features contributing to that understanding?

Outline

- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Hypotheses

1. Participants who have greater levels of training, across all modulation types, will be more accurate. (1, 2)
2. Responses to the modulations will depend on the modulation type. (3, 4)
   a. Responses to direct modulations will be the most accurate
   b. Responses to the common tone modulations will be next most accurate
   c. Responses to the pivot chord modulations will be least accurate
3. Key distance and mode change will be more accurate predictors of modulation perception. (3, 4)
4. Trained listeners will respond faster to the modulations than untrained listeners. (2)

Outline

- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Procedure

- Informed consent obtained
- Music questionnaire survey
- Explanation of experiment:
  - Listening for modulations
  - Respond as many times as you like
  - Respond as soon as you think the music has moved to a new key area.
- Experiment
- Informal debrief, answering any questions.
- Analysis:
  - 3x3 mixed ANOVA: A' for participants, by training level & modulation type
  - Between Groups ANOVA for excerpts on A', by mode change
  - Regression analyses for excerpts using key distance and mode change as independent variables
  - 3 way between groups ANOVA: response time for participants

Participants

- Recruitment
  - Majority from SONA
  - Music department at NSULA
  - Professional musicians & music educators in the DFW area and the North Texas/I-20 corridor between DFW & Shreveport
- 180 participants (M = 92, F = 87, NB = 1)
  - Aged 18 – 59 (M = 22.9, SD = 5.49)
- Screened for:
  - Exposure to or training in Carnātic Music
  - Absolute pitch
  - Hearing disability (deafness, tinnitus, or amusia)

Group assignments

- Three groups based on level of music training
  - Untrained/Non-musicians: 0 – 2 years of music training (n = 60, M = 0.63, SD = 0.92)
  - Moderately trained: 3 – 9 years of music training (n = 60, M = 5.53, SD = 1.75)
  - Highly trained: 10+ years of music training (n =60, M = 16.07, SD = 7.75)
    - OR had < 10 years formal training but had successfully completed an AP™ or university – level ear-training/music theory course. (n = 4)

Stimuli

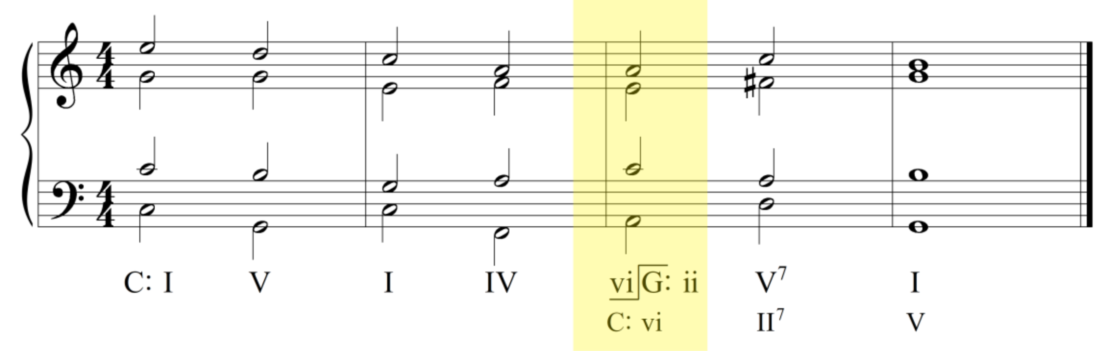
- 49 total excerpts by Classical and Romantic composers
  - Composition dates between 1762 – 1890
  - Featuring the works of Joseph Haydn, Roman Hofstetter, Wolfgang Amadeus Mozart, Ludwig van Beethoven, Franz Schubert, and Johannes Brahms
- 14 featuring each type of modulation + seven non-modulating excerpts
- Selection criteria & balancing
- Total listening of time of 22m 59s.
- Ripped from an audio CD using fre:ac & presented using the .wav file format to ensure presentation quality.

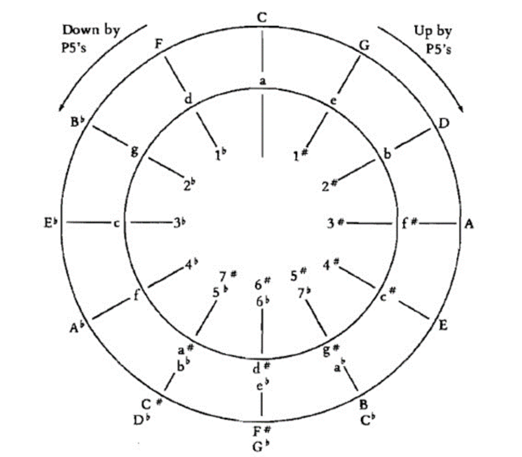
Stimuli

- Three types of modulations
  - Pivot Chord
  - Direct
  - Common Tone

Pivot Chord Modulation

- Smooth transition
- Usually modulates to V, or another close key
- Requires a chord that is common to both the starting and target keys.





Benjamin, Horvit, and Nelson (2003)

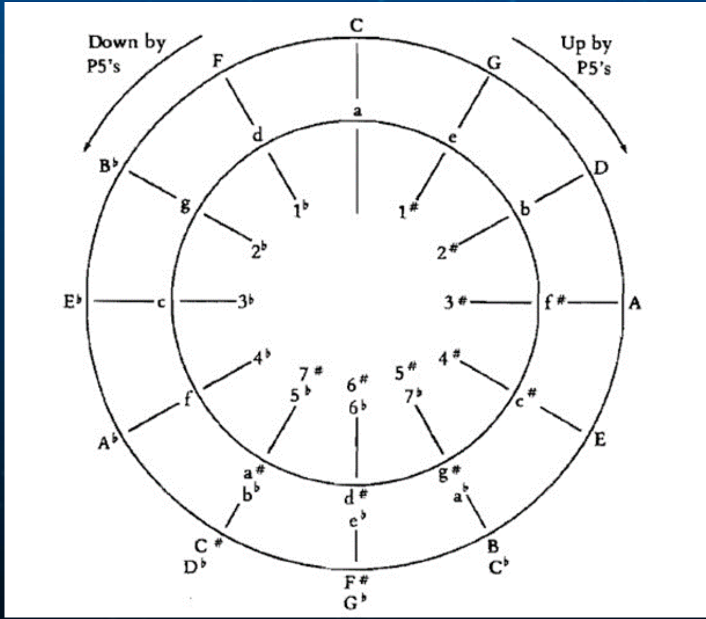
Movement I, Allegro Molto, ms. 1-27

School of Brain and Behavioral Sciences

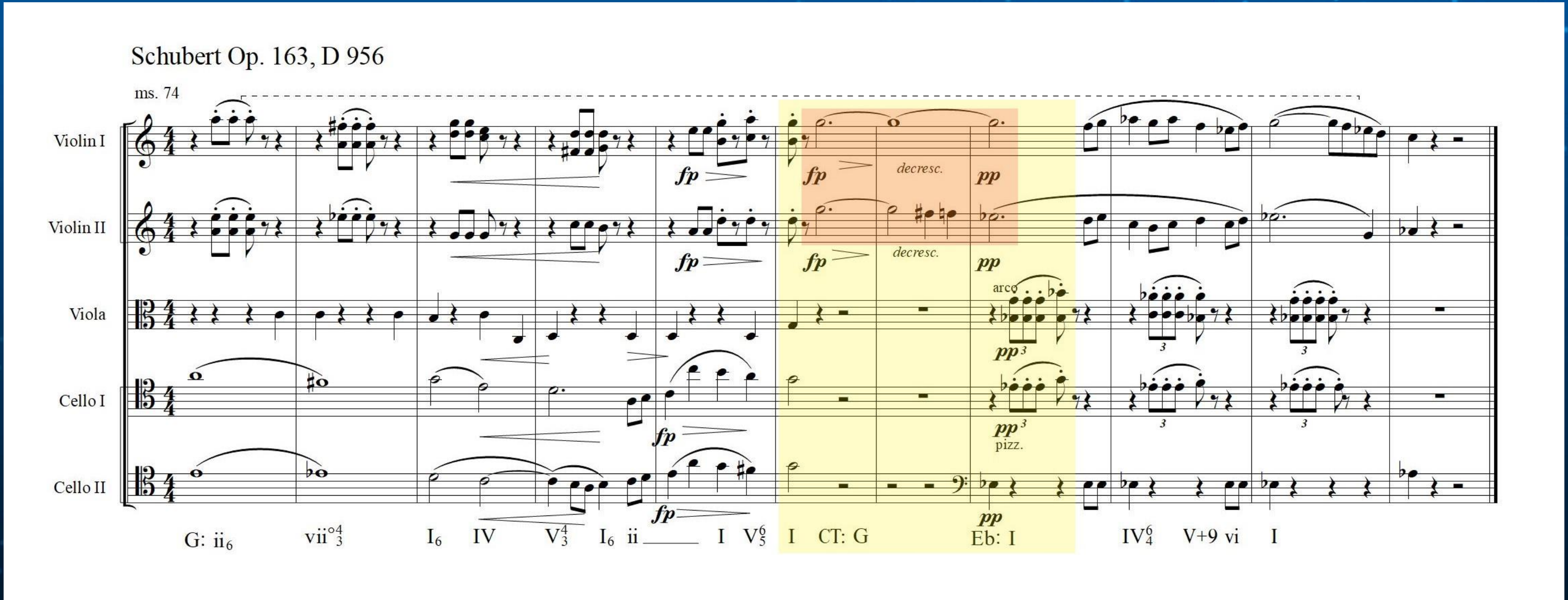
THE UNIVERSITY OF TEXAS AT DALLAS

Common Tone Modulation

- Smooth transition
- Requires a common tone between starting and target keys.
- Usually used to modulate to a distant key, often a tonic interval of a third between starting and target keys.



Benjamin, Horvit, and Nelson (2003)

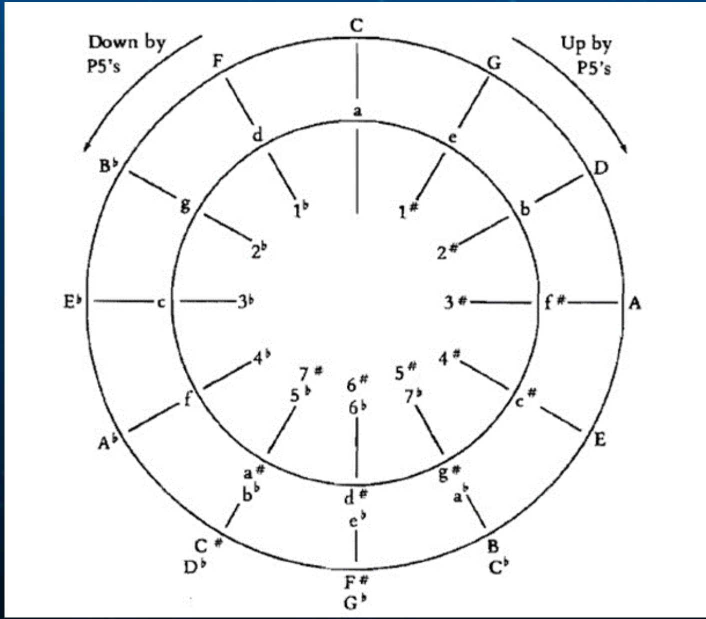


School of Brain and Behavioral Sciences

THE UNIVERSITY OF TEXAS AT DALLAS

Direct Modulation

- Can seem abrupt or jarring.
- Immediate change between keys
- No common tones necessary
- Can modulate to a near or distant key, often the dominant or the submediant (relative minor)

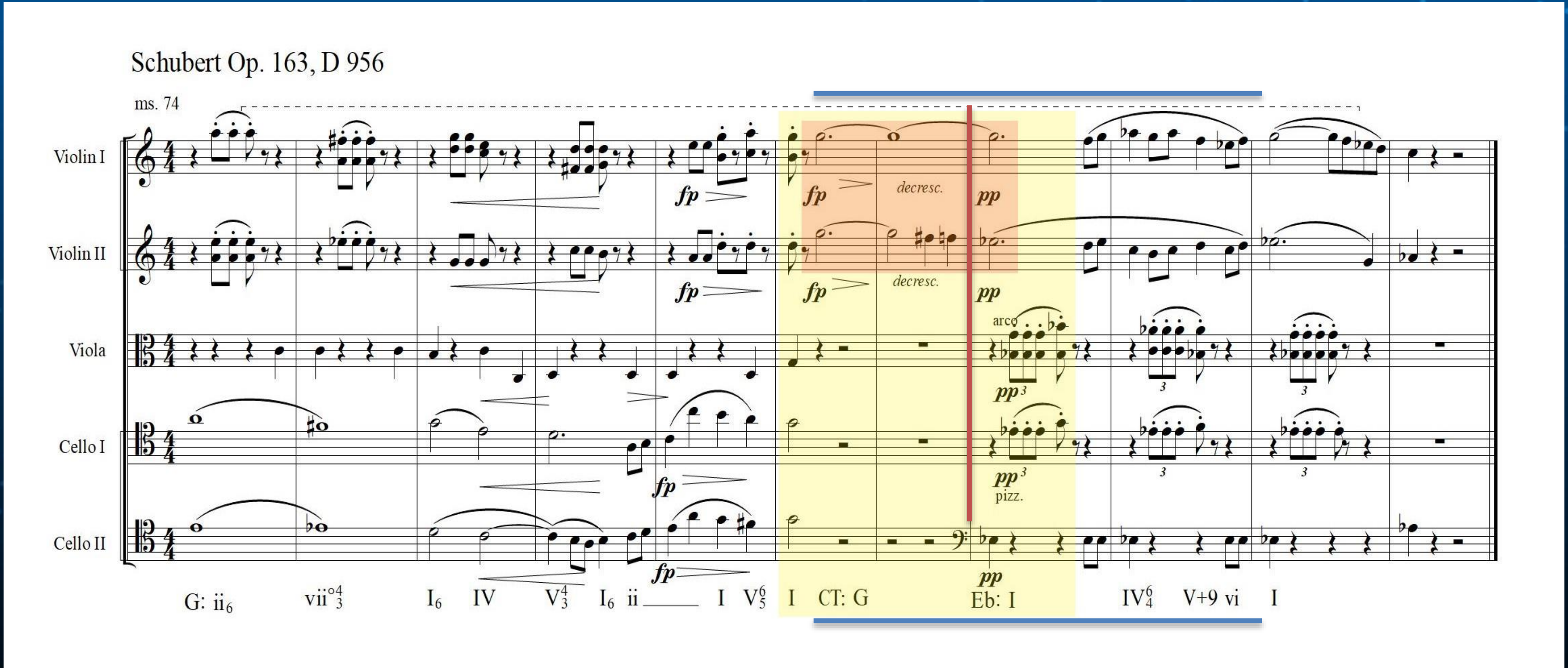


Benjamin, Horvit, and Nelson (2003)

School of Brain and Behavioral Sciences

THE UNIVERSITY OF TEXAS AT DALLAS

Presentation



Outline

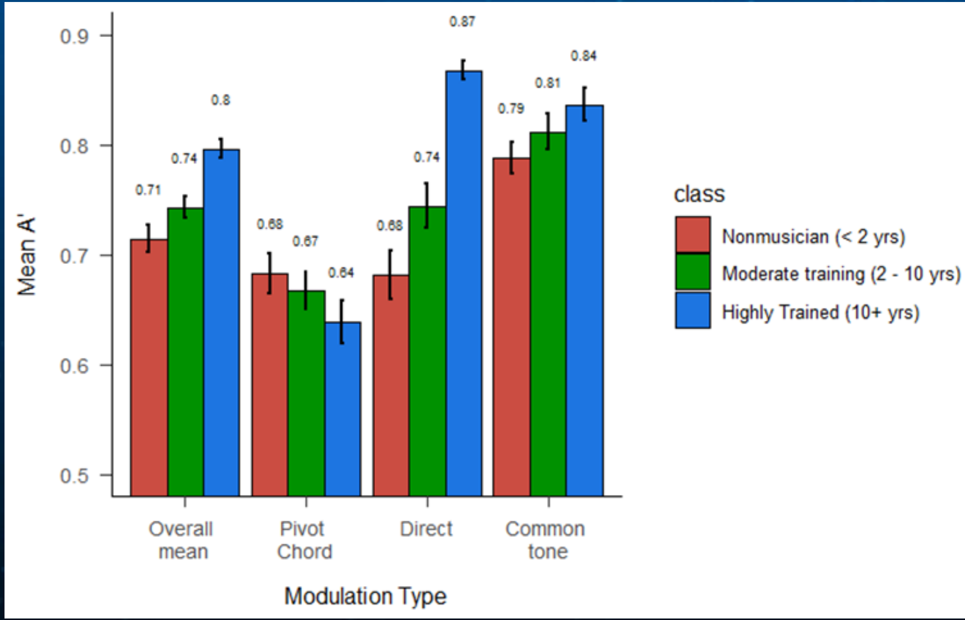
- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

Results: Training & Modulation Type

Simple Training:
$F(2,531) = 10.51,\ MSE = 0.01, p < .001$

Simple Modulation type:
$F(2,531) = 59.48, MSE = 0.02,\ p < .01$

Interaction:
$F(4, 531) = 11.73, MSE = 0.02, p < .001.$



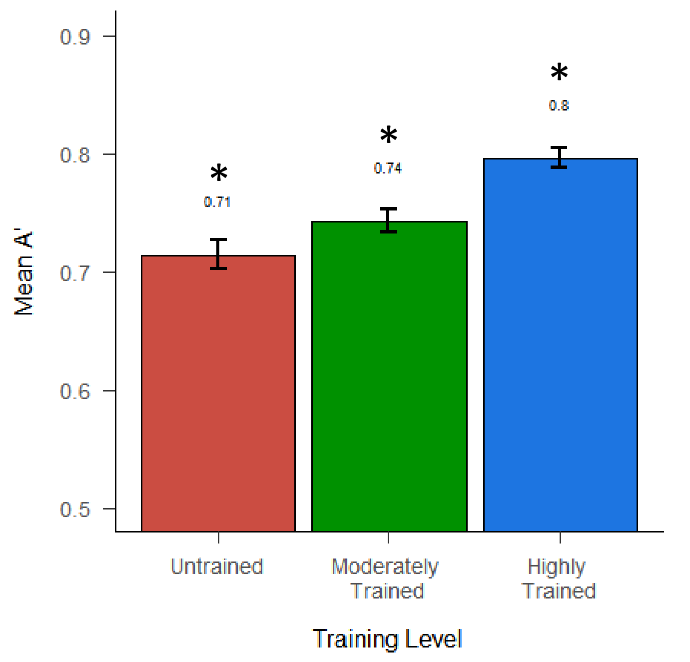
Results: Training & Modulation Type

Simple Training:

Highly Trained vs. Untrained:
d = -0.41, 95% CI [.21, 0.62], $p < .001$

Highly Trained vs. Moderately Trained:
d = -0.26, 95% CI [0.05, 0.46], $p = .01$

Moderately Trained vs. Untrained:
d = -0.15, 95% CI [-0.05, 0.35], $p = .22$ (NS)



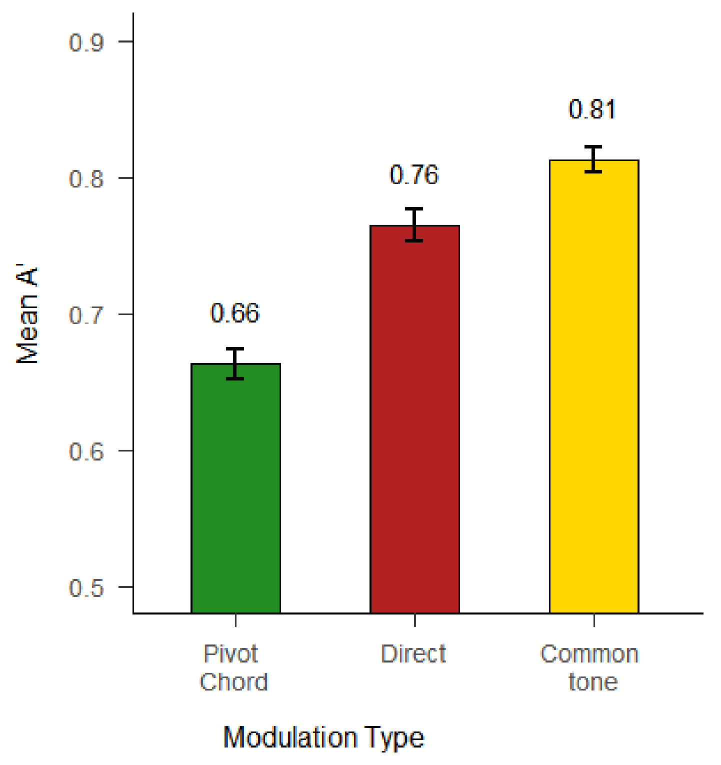
Results: Training & Modulation Type

<u>Simple Modulation type:</u>

Direct – Pivot Chord:
d = -0.66, 95% CI [0.45, 0.88], *p* < .001

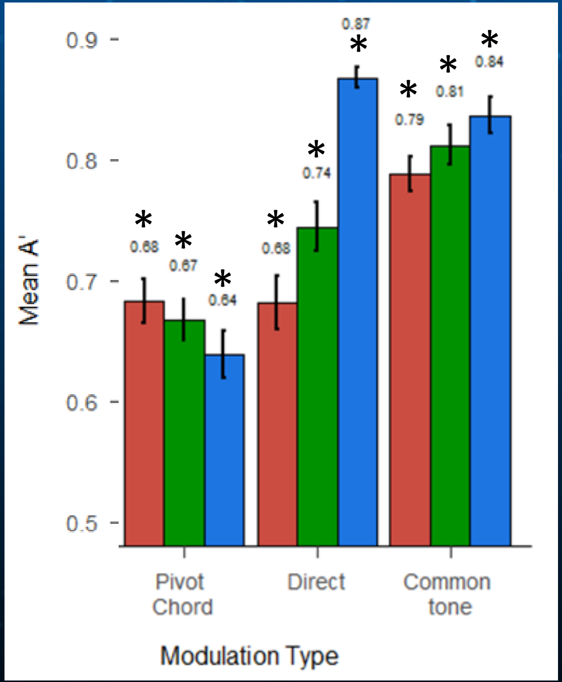
Common Tone – Pivot Chord:
d = -0.97, 95% CI [0.77, 1.19], *p* < .001

Common Tone – Direct:
d = -0.31, 95% CI [0.10, 0.52], *p* = .002



Results: Training & Modulation Type

	Cohen's *d*	lower limit	upper limit	p value
1 - CT vs. 1 - PC	0.69	0.48	0.90	.001
1 - CT vs. 1 - DM	0.69	0.48	0.91	< .001
2 - DM vs. 2 - PC	0.50	0.29	0.71	.041
2 - CT vs. 2 - PC	0.95	0.73	1.16	< .001
3 - DM vs. 3 - PC	1.49	1.26	1.72	< .001
3 - CT vs. 3 - PC	1.29	1.06	0.91	< .001



Results: Key distance & mode change

Simple Mode Change

$F(1,40) = 0.04$, $MSE = 0.02$, $p = .84$ (ns)

Simple Key Distance

$R^2_{adj} = .12$, 90% CI[0.02, 0.33], $F(1,40) = 6.25$, $p = .02$

$b = -.15$, 95% CI[-0.27, -0.03], $t(40) = -2.55$, $p = .02$

Key distance by Mode Change

$R^2_{adj} = .17$, 90% CI[0.03, 0.39], $F(3, 38) = 3.86$, $p = .02$

b (key distance) = -0.04, 95% CI[-0.22, 0.13],

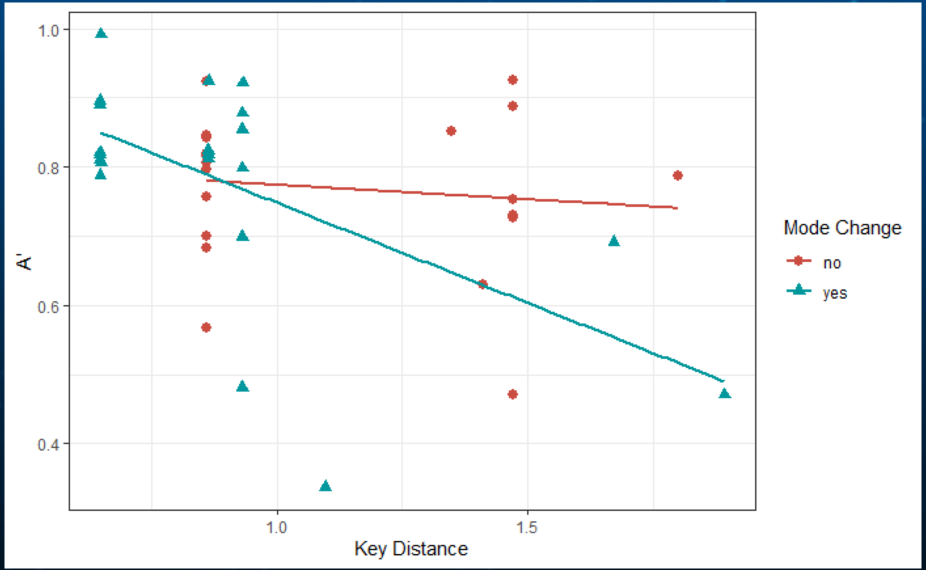
$t(38) = -0.49$, $p = 0.63$

b (mode change) = 0.22, 95% CI[-0.04, 0.49],

t(38) = 1.71, $p = 0.096$

b (interaction) = -0.25, 95% CI[-0.49, 0.00],

t(38) = -2.03, p = 0.49



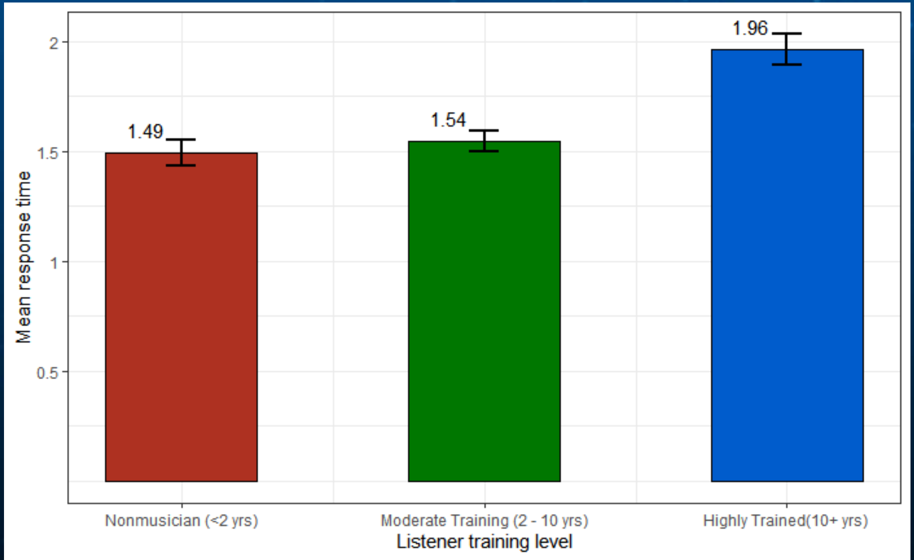
y = 0.82 – 0.04a + 0.22b – 0.25ab

Results: Reaction Time

Simple Training:
$F(2,177) = 19$, $MSE = 0.21$, $p < .001$

Simple Modulation type:
$F(2,39) = 1.71$, $MSE = 0.58$, $p = .194$ (ns)

	Cohen's *d*	lower limit	upper limit	*p* value
Untrained vs. Moderate training	0.07	-0.14	0.25	.79 (ns)
Untrained vs. Highly trained	0.64	0.27	0.67	<.001
Moderate training Vs. Highly trained	0.57	0.22	0.62	<.001



Outline

- Background
- Questions
- Hypotheses
- Methods
- Results
- Discussion

- Hypothesis 1:
  - Participants who have greater levels of training, across all modulation types, will be more accurate.
    - Overall means support this hypothesis
    - Untrained listeners performed above chance across modulation types
    - Pivot chord results contradict this hypothesis

- Hypothesis 2:
  - Response accuracy will depend on the modulation type.
    1. Responses to direct modulations will be the most accurate
    2. Responses to the common tone modulations will be next most accurate
    3. Responses to the pivot chord modulations will be least accurate

  - Actual:
    1. Most accurate: Common Tone (A’ = .81)
    2. Middle: Direct (A’ = .76)
    3. Least accurate: Pivot Chord (A’ = .66)

- Hypothesis 3:
  - Key distance and mode change will be more accurate predictors of modulation perception. (1, 3)
  - Not supported.
    - Participants were actually less accurate given greater key distance.
    - Mode change exacerbated this effect.
  - Possibly confounded by the number of excerpts that modulated to distant keys.

- Hypothesis 4:
  - Trained listeners will respond faster to the modulations than untrained listeners.
  - Not supported. Trained listeners reacted more slowly than either of the other groups.

Conclusions

1. Listeners, across training levels, track tonic region independent of surface features.
2. Training helps, but only when that training is at or approaches a professional level.
3. The most helpful surface feature is a sustained pitch that both provides reference and time to allow for listener comprehension.
4. Trained listeners take longer to respond, but are overall more accurate.
5. Prior evidence regarding key distance and modulation perception, specifically cognitive lag in processing greater key distance, is supported.
6. Highly trained listeners seem to be able to consciously access the information regarding pitch set content and the specific function of each pitch in the set.

Limitations & Future Directions

- Possible limitations that should be addressed:
  - Selecting more excerpts with greater key distance.
  - Better account for phrase boundary in stimuli creation to rule out any specific effects of phrase boundary.
  - Harmonic language & complexity can be different between compositional styles, balancing in this regard could rule out the effect of period.

- Future directions:
  - Cross cultural studies using other musical idioms and cultures
  - Analyses featuring age and passive exposure to music
  - More research into the cognitive lag question brought up by the results of the timing experiment and trained listener's results on the

Thanks!

- UTDallas MPaC
  - Dr. Dowling
  - Dr. Raman
  - Kieth Gryder
  - Cynthia Chan
- BBS Faculty & Students
- NSULA CAPA Faculty & Staff
  - Dr. Adam Hudlow
  - Dr. Mitch Davis
- Mr. Rance Hawthorne