

# Statistical Analysis of Timbre Descriptors on Musical Instrument Classification

---

Song Hui Chon

MUMT 621

9 April 2009

# Overview

---

- Motivation
- Research Questions & Current Project
- Analyses & Result
- Conclusion

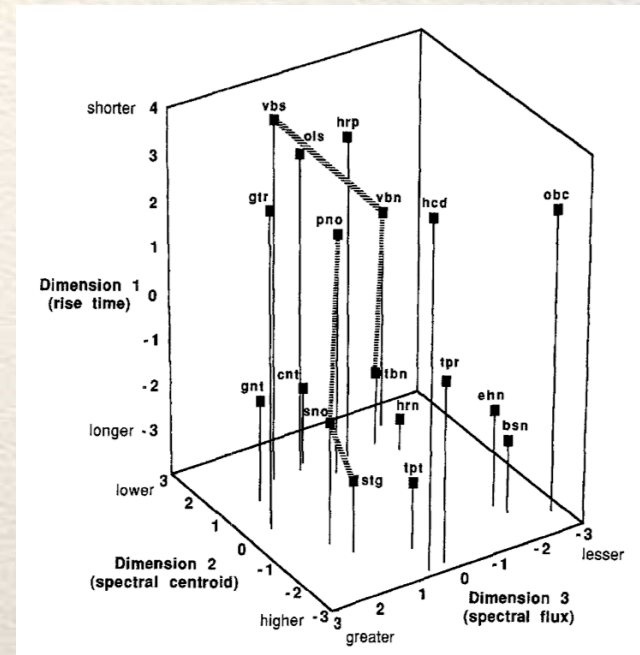
# Motivation

---

- Musical instrument classification
  - Features based on information on pitch, timbre, excitation method, etc.
  - Want to study timbre aspect
- Timbre descriptors
  - Over 70 descriptors in research [Rioux et al. 2006]
  - Little systematic evaluation done so far

# Timbre

- "[...] that attribute of sensation in terms of which a listener can judge that two sounds having the same loudness and pitch are dissimilar" [ANSI 1973]
- Multidimensional perception of sound
  - Many attempts to model *timbre space* [Grey 1977] [Grey & Gordon 1978] [Krumhansl 1989] [Iverson & Krumhansl 1993] [McAdams et al. 1995] [Lakatos 2000]
  - No unified model developed yet



# Musical Instrument Classification

---

- One of major MIR research areas
- Considers many features, including some timbre related ones
- Usually employs machine-learning algorithms for better performance

# Research Questions

- What are the relationships in timbre descriptors?
- What are the best and the worst descriptors in instrument family classification?
- Are there certain stimuli that are consistently worse than others in instrument family classification?
- What are the best and the worst descriptors in excitation classification in wind instrument family?

# Current Project

- A statistical approach
  - Performance comparison of 70 timbre-related descriptors
  - Feature values computed on 135 stimuli from well-known timbre studies ([Grey 1977] [Grey & Gordon 1978] [Krumhansl 1989] [Iverson & Krumhansl 1993] [McAdams et al. 1995] [Lakatos 2000])
  - Analysis on 114 stimuli, after excluding those that are hybrids or have not been analyzed before

[Rioux et al. 2006]

# Timbre Descriptors

GROUP	DESCRIPTORS
1. Spectral slope	DECI, DECIDB, DECS
2. Spectral centroid	CGSILO, CGSMAX, CGSRMS, CGSMOY, CGSH, CGSI, CGSB, CGSA, CGS, CGSB2, CGSC
3. Spectral flux	FLMAX, FLRMS, FLMOY, FLI
4. Spectral spread (STD)	STDILO, STDI, STDH, STDMOY, STDRMS, STDMAX, STDB, STD, STDB2, STDC, STDA
5. Spectral deviation	DEVMAX, DEVRMS, DEVMOY, DEVIDB, DEVS
6. Spectral shape	SKEW, KURT, IPH, VSRATE, SLOPE, STDIDB, CGSIDB
7. Fluctuation & Roughness	MAGCO, ROUGH, FLUC, MAXIMUM
8. RMS power & energy	LDBA, LDBB, LDBC, LDB, NRGB, NRGBI, NRGH CGT, STDT, ED, ITMPN1, MIX, ACUM
9. Attack time	DEVI, HAC, ITMPN3, ITMPN2, ITMPN4, VSPH, VSPC, LTMM, LTMLM, LTMR, LTMLR, LAT



# Instruments

Family	Instruments
STRING	<p>[Grey 1977] [Grey &amp; Gordon 1978] “string”</p> <p>[Iverson &amp; Krumhansl 1993] Cello, Piano, Violin</p> <p>[Lakatos 2000] Harp, Harpsichord, Piano, Violin (“martele”, “no vibe”)</p> <p>[Krumhansl 1989] [McAdams et al. 1995] Guitar, Harpsichord, Harp, Piano (generic, bowed, sampled), “string”</p>
WIND	<p>[Grey 1977] [Grey &amp; Gordon 1978] Bassoon, Clarinet, English Horn, French Horn, Flute, Oboe, Trombone, Trumpet, Saxophone</p> <p>[Iverson &amp; Krumhansl 1993] Bassoon, Clarinet, English Horn, Flute, French Horn, Oboe, Saxophone, Trombone, (Muted) Trumpet, Tuba</p> <p>[Lakatos 2000] Clarinet, Crumhorn, English Horn, Flute (flutter, no vibe), French Horn, Organ, Baroque Recorder, Saxophone, (Muted) Trumpet</p> <p>[Krumhansl 1989] [McAdams et al. 1995] Bassoon, Clarinet, English Horn, French Horn, Oboe, Trombone, Trumpet</p>
PERCUSSION	<p>[Iverson &amp; Krumhansl 1993] Tubular Bells, Vibraphone</p> <p>[Lakatos 2000] Bamboo Chimes, Bongo, Castanets, Celesta, Cuica, Cymbal (bowed, struck), Log drum, Marimba, Snare, Steel drum, Tam Tam, Tambourine, Temple Block, Tubular Bells, Tympani, Vibraphone (bowed, hard mallet)</p> <p>[Krumhansl 1989] [McAdams et al. 1995] Vibraphone</p>

# Analysis

---

- Correlation Analysis
- K-means Clustering
- Adjusted rand index
  
- MATLAB and SPSS used

# Question 1

---

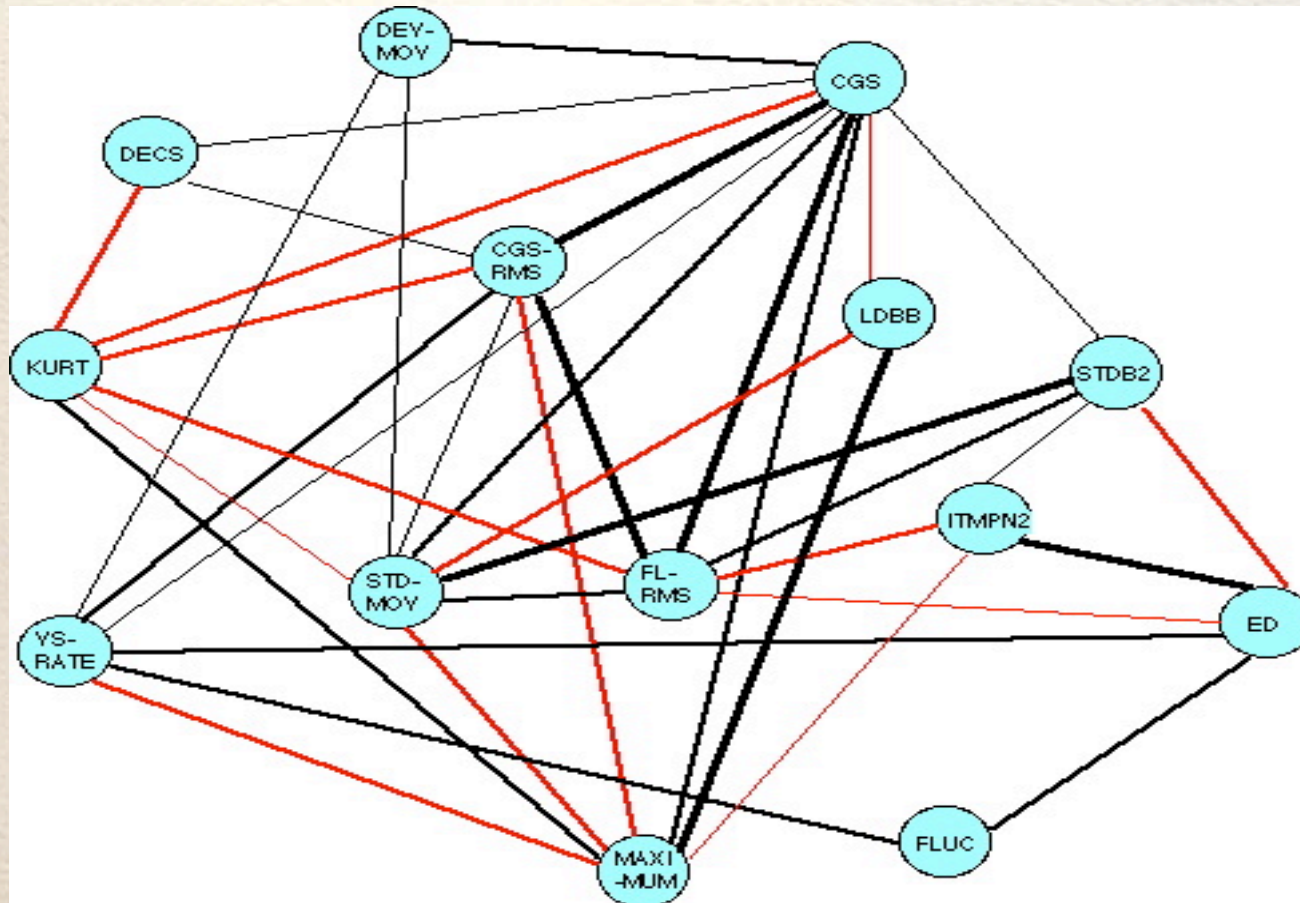
- What are the relationships in timbre descriptors?
  - Correlation Analysis

# Correlation Analysis

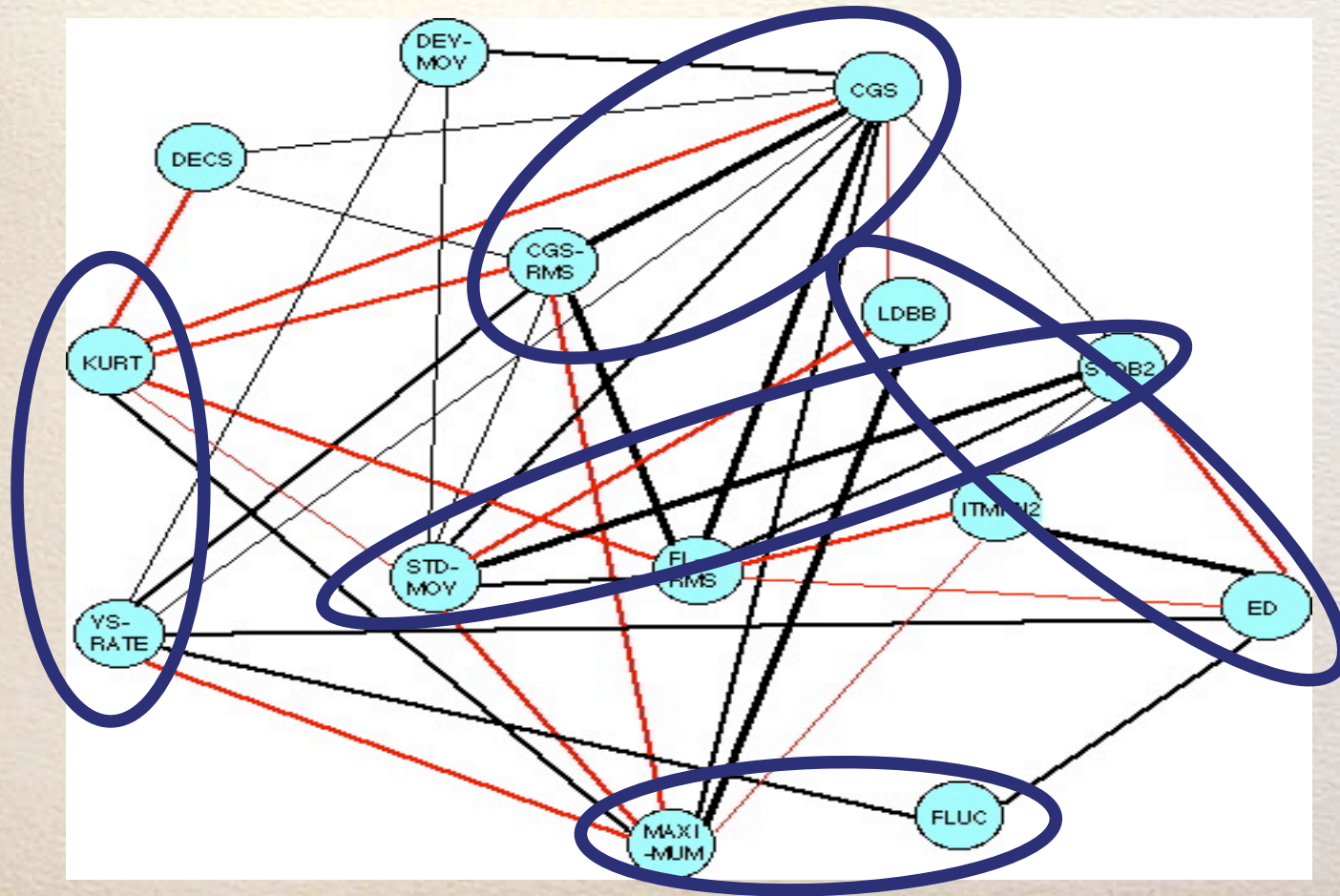
- 9 Groups [Rioux et al. 2006]

1. Spectral slope	DECI, DECIDB, <b>DECS</b>
2. Spectral centroid	CGSILO, CGSMAX, <b>CGSRMS</b> , CGSMOY, CGSH, CGSI, CGSB, CGSA, <b>CGS</b> , CGSB2, CGSC
3. Spectral flux	FLMAX, <b>FLRMS</b> , FLMOY, FLI
4. Spectral spread (STD)	STDILO, STDI, STDH, <b>STDMOY</b> , STDRMS, STDMAX, STDB, STD, <b>STDB2</b> , STDC, STDA
5. Spectral deviation	DEVMAX, DEVRMS, <b>DEVMOY</b> , DEVIDB, DEVS
6. Spectral shape	SKEW, <b>KURT</b> , IPH, <b>VSRATE</b> , SLOPE, STDIDB, CGSIDB
7. Fluctuation & Roughness	MAGCO, ROUGH, <b>FLUC</b> , <b>MAXIMUM</b>
8. RMS power & energy	LDBA, <b>LDBB</b> , LDBC, LDB, NRGB, NRG1, NRGH CGT, STDT, <b>ED</b> , ITMPN1, MIX, ACUM
9. Attack time	DEVI, HAC, ITMPN3, <b>ITMPN2</b> , ITMPN4, VSPH, VSPC, LTMM, LTMLM, LTMR, LTMLR, LAT

# Correlation Analysis



# Correlation Analysis

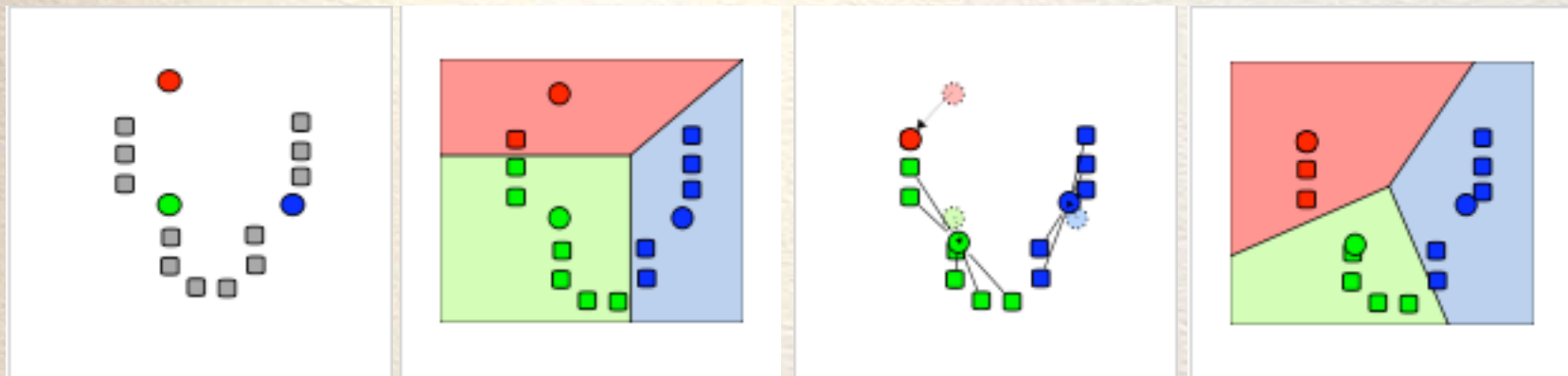


## Question 2

---

- What are the best and the worst descriptors in instrument family classification?
  - K-means Clustering
  - Adjusted rand index

# K-means Clustering



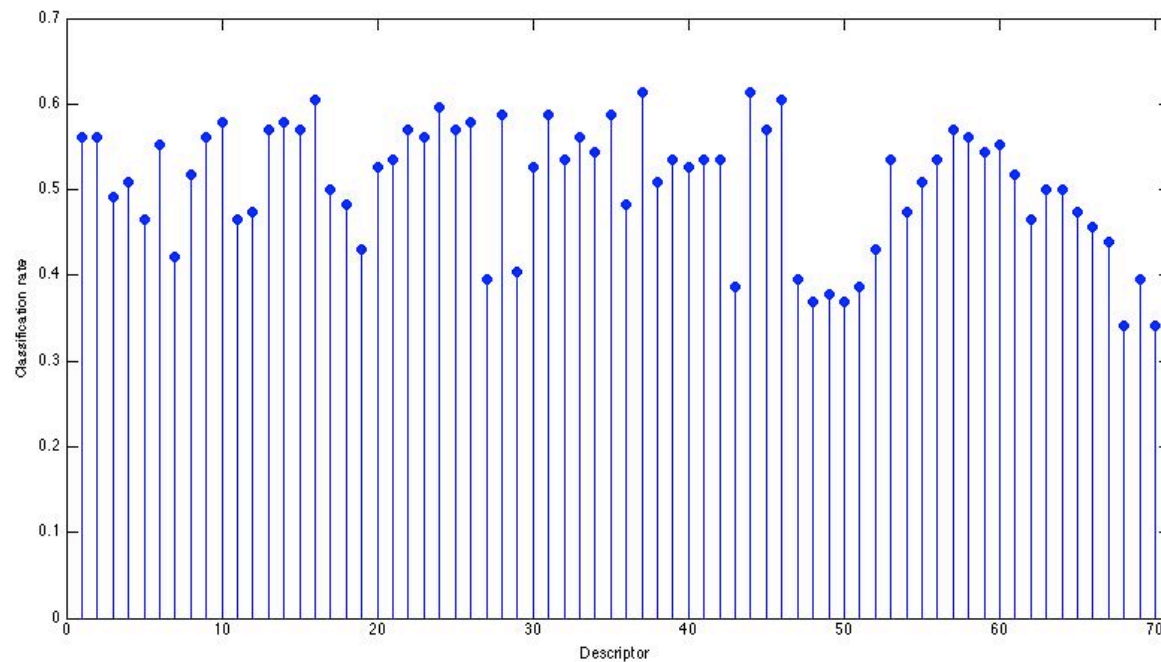
[Wikipedia K-means Algorithm]



# Instrument Family Classification

- K-means Clustering Result

[min = .3421    mean = .5073    max = .6140]





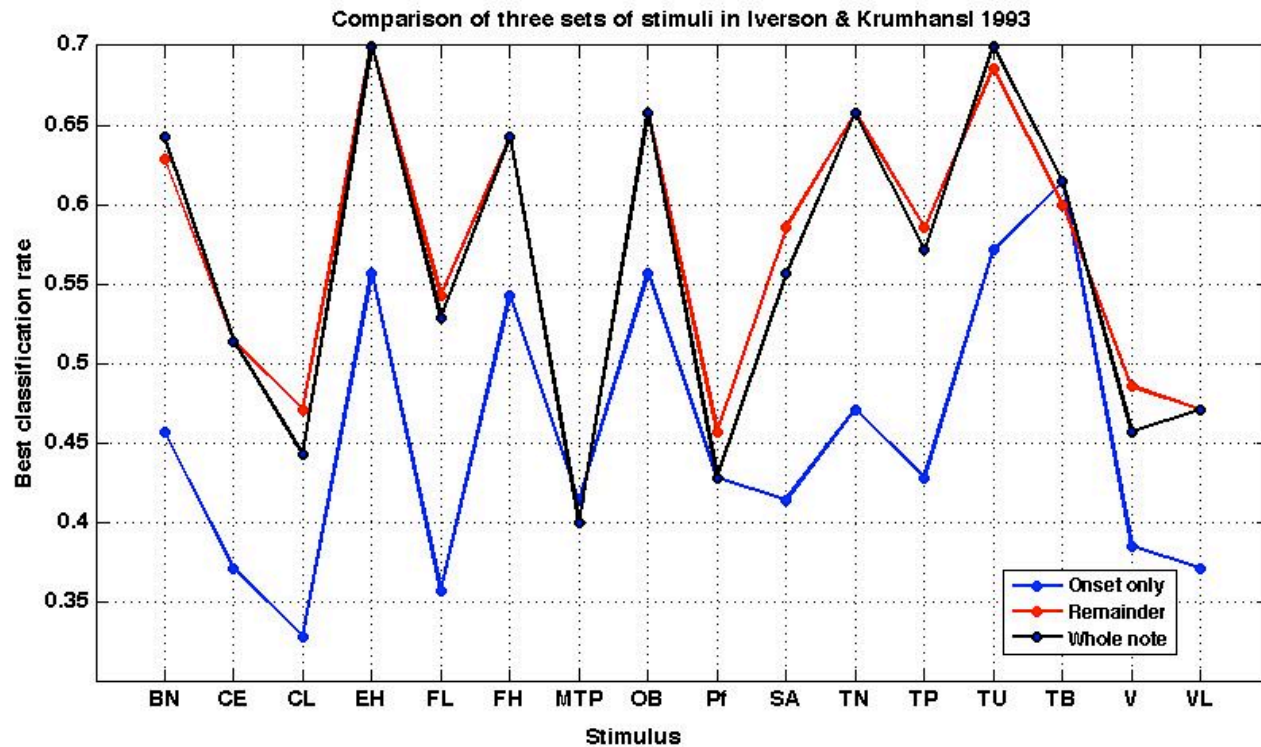
## Question 3

---

- Are there certain stimuli that are consistently worse than others in instrument family classification?
  - K-means Clustering

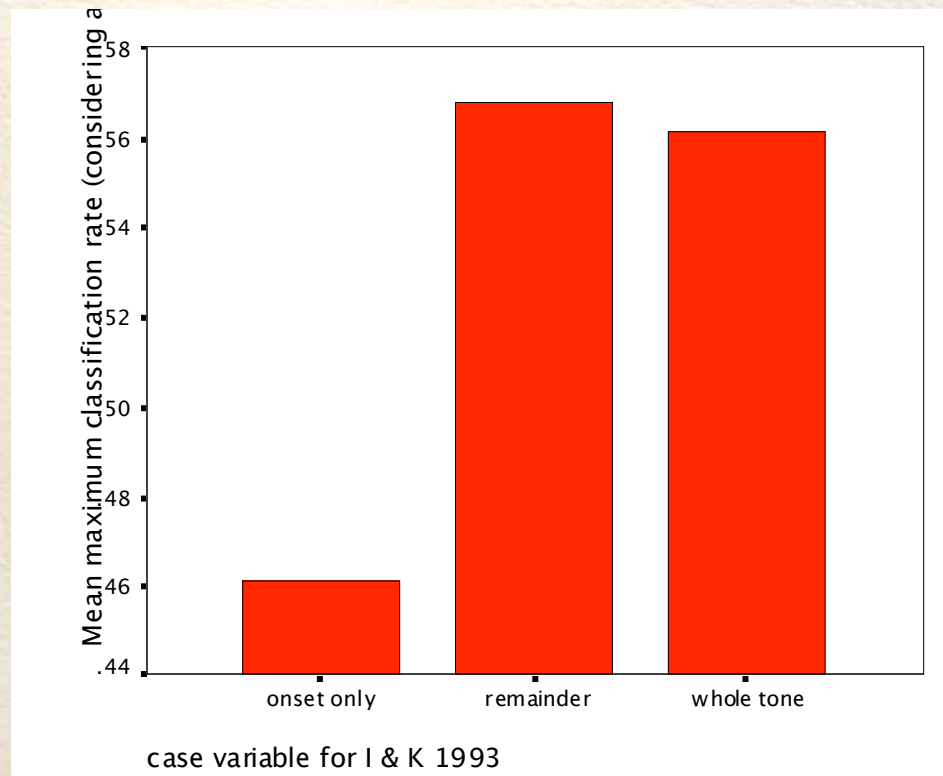
# Onset, Remainder or Whole Tone?

- Iverson & Krumhansl 1993



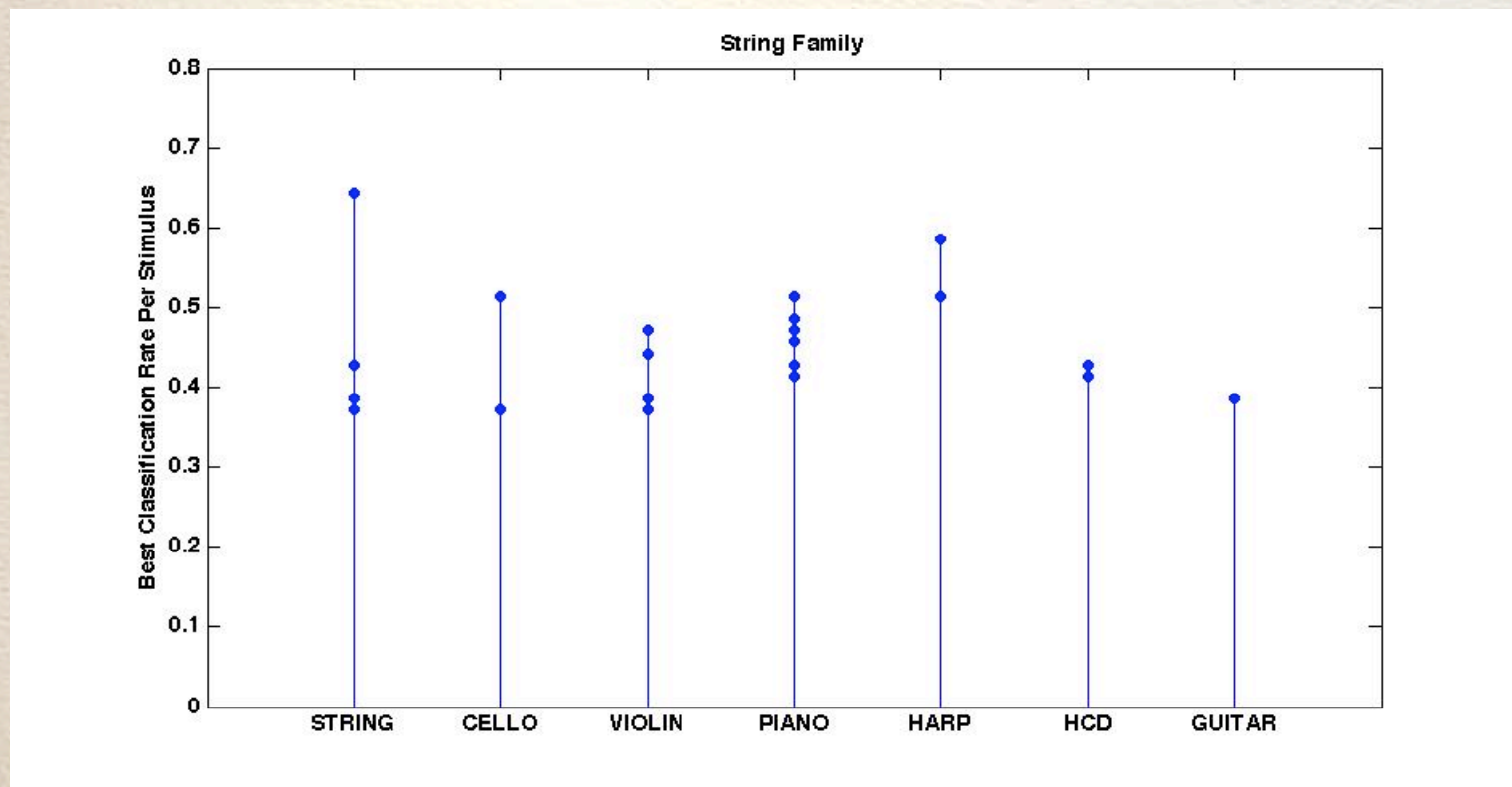
# Onset, Remainder or Whole Tone?

- Iverson & Krumhansl 1993



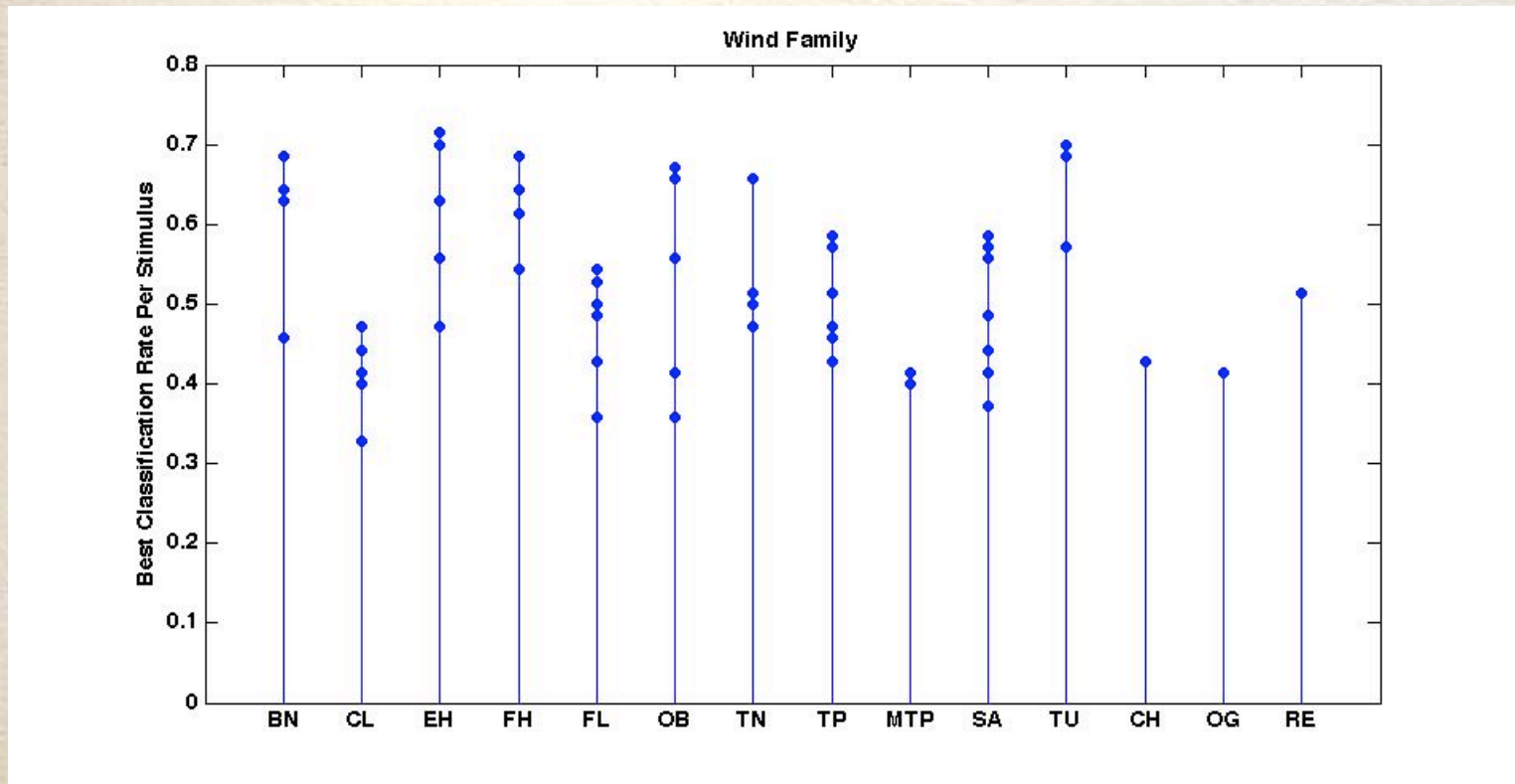
# Worse Stimuli?

- String Family



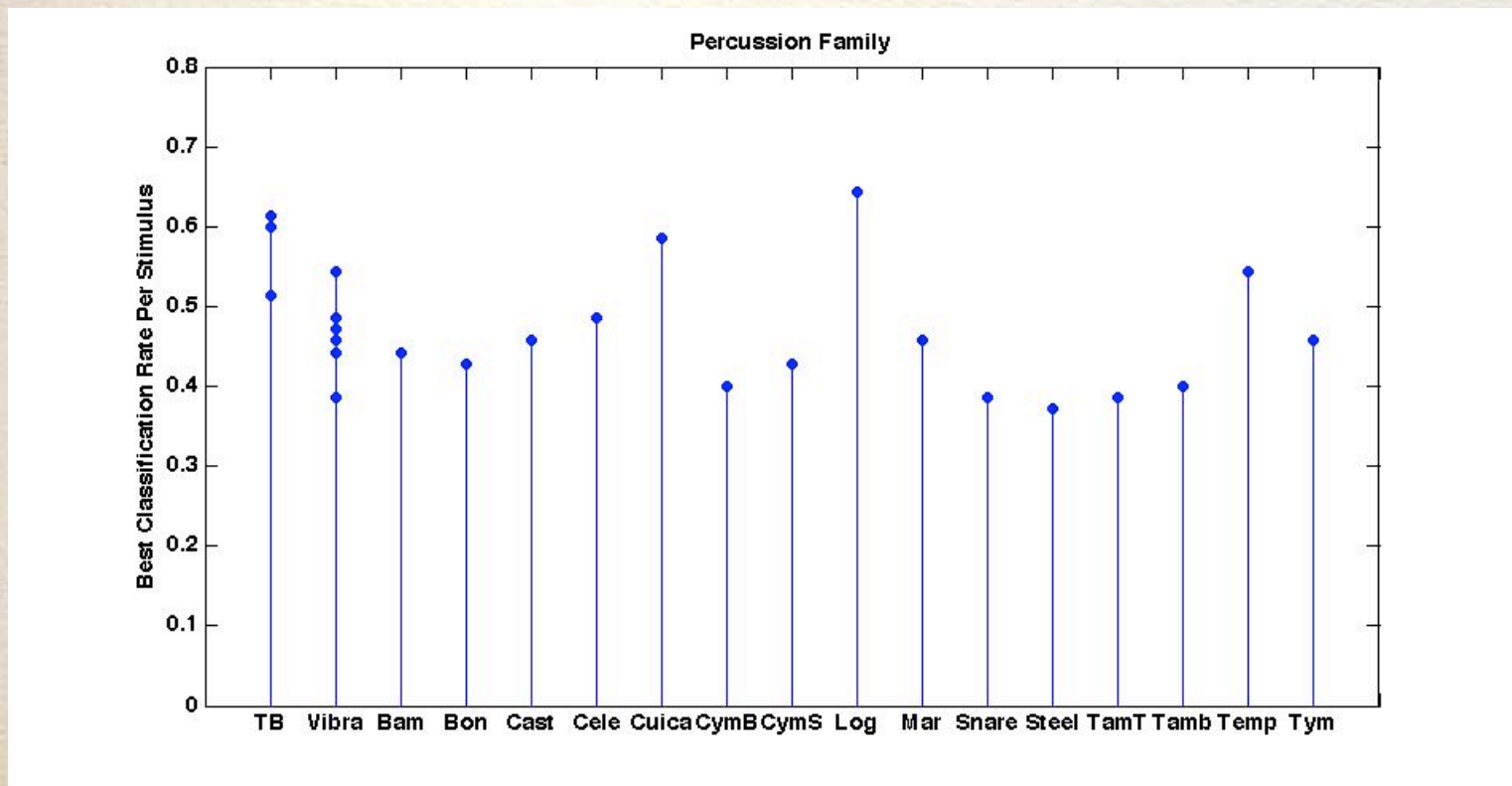
# Worse Stimuli?

- Wind Family



# Worse Stimuli?

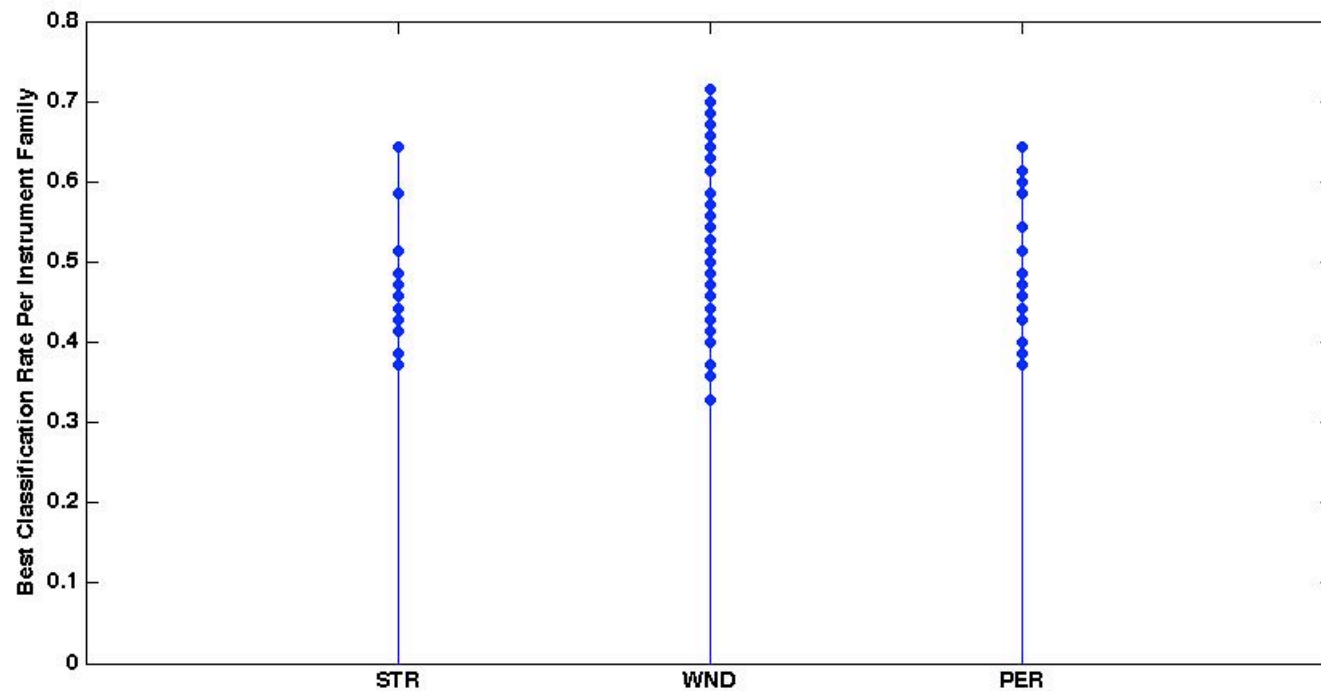
- Percussion Family





# Worse Stimuli?

- All together



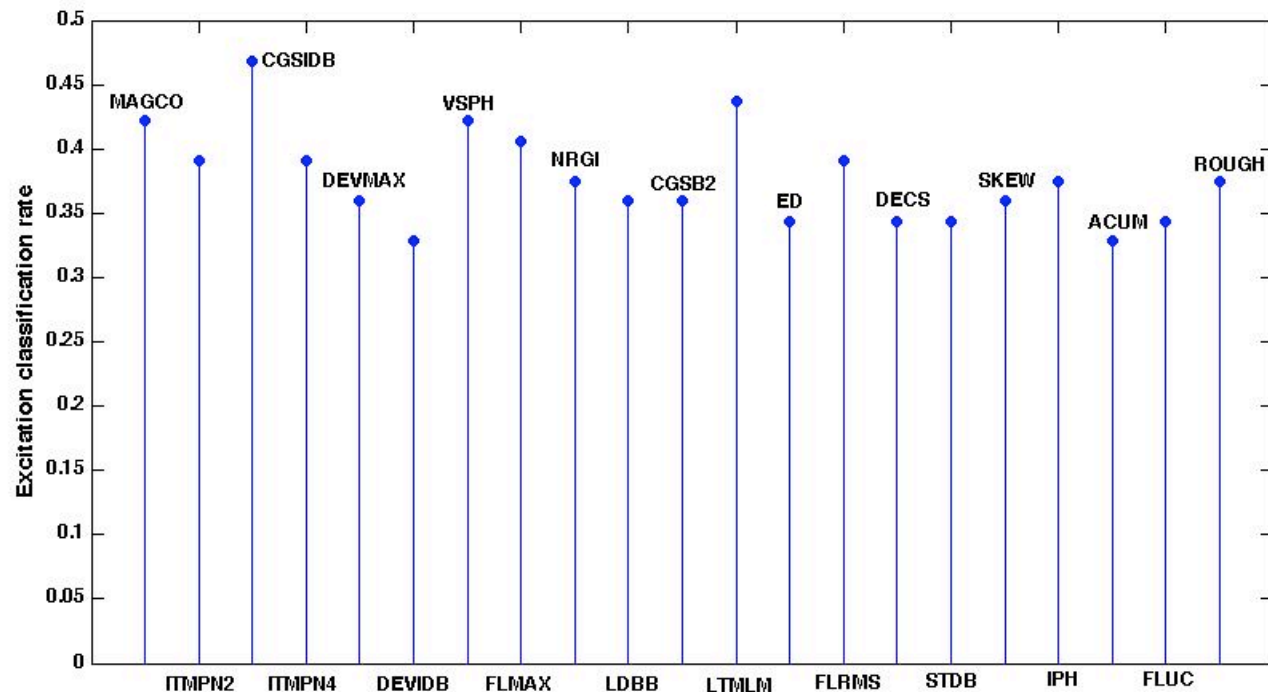
# Question 4

- What are the best and the worst descriptors in excitation classification in wind instrument family?
  - K-means Clustering
  - Adjusted rand index
  - Wind instrument family (64 stimuli)
    - 4 types of excitation

Excitation	Instruments
Airjet	Flute, Baroque Recorder
Single Reed	Clarinet, Saxophone
Double Reed	Oboe, Bassoon, Crumhorn, English Horn
Lip Reed	(Muted) Trumpet, Trombone, French Horn, Tuba

# Wind Instrument Excitation Classification

- 21 selected descriptors



# Conclusion

---

- Statistical analysis of 70 timbre descriptors
  - Confirmed 9 clusters from previous study using Correlation Analysis
  - K-means clustering was used for instrument family classification and excitation type classification in wind instruments.
    - Not optimal clustering
    - Some descriptors better and more stable than others
    - Some instruments harder to classify than others

# Future Works

---

- Consider combining descriptors for better classification
- Discriminant analysis for optimal linear classification
- More data!
- Are there any acoustical or psychoacoustical explanations on certain instruments being easier or harder to classify?
- Bridge with timbre space model

# References

- [ANSI 1973] American National Standard Institute. 1973. American national psychoacoustical terminology. S3.20. New York: American Standards Association.
- [Borg & Groenen 1997] Borg, I., & Groenen, P. 1997. *Modern multidimensional scaling*. New York, NY: Springer-Verlag.
- [Caclin et al. 2005] Caclin, A., McAdams, S., Smith, B.K., & Winsberg, S. 2005. Acoustic correlates of timbre space dimensions: a confirmatory study using synthetic tones. *Journal of Acoustical society of America*, 118(1), 471—82.
- [Clark et al. 1964] Clark, M., Robertson, P., & Luce, D.A. 1964 A preliminary experiment on the perceptual basis for musical instrument families. *Journal of the Audio Engineering Society*, 12, 199—203.
- [Grey 1977] Grey, J.M. 1977. Multidimensional perceptual scaling of musical timbres. *Journal of Acoustical Society of America*, 61, 1270—1277.
- [Grey & Gordon 1978] Grey, J. M., & Gordon, J. W. 1978. Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, 63, 1493—1500.
- [Giordano & McAdams 2009] Giordano, B., & McAdams, S. 2009. Mechanics of the sound source and the perception of musical timbre: Evidence from previous studies. *Submitted*.
- [Handel & Erickson 2004] Handel, S., & Erickson, M.L. 2004. Sound source identification: the possible role of timbre transformations. *Music Perception*, 21(4), 587—610.

# References

- [Helmholtz 1877] Helmholtz, H.L.F. 1877. *On the Sensations of Tone as a Psychological Basis for the Theory of Music*. New York, NY: Dover (English translation in 1954).
- [Herrera-Boyer et al. 2003] Herrera-Boyer, P., Peeters, G., Dubnov, S. 2003. Automatic Classification of Musical Instrument Sounds. *Journal of New Music Research*, Vol. 32, No. 1, 3—21.
- [Iverson & Krumhansl 1993] Iverson, P., & Krumhansl, C.L. 1993. Isolating the dynamic attributes of musical timbre. *Journal of Acoustical Society of America*, 94, 2593—2603.
- [Krumhansl 1989] Krumhansl, C.L. Why is musical timbre so hard to understand? In *Structure and Perception of Electroacoustics Sound and Music*, edited by S. Nielzen and O. Olsson (Excerpta Medica, Amsterdam), 44—53.
- [Krimphoff et al. 1994] Krimphoff, J., McAdams, S., & Windberg, S. 1994. Caractérisation du timbre des sons complexes. II: Analyses acoustiques et quantification psychophysique. *Journal de Physique*, 4, 625—628.
- [Krumhansl 1989] Krumhansl, C.L. 1989. Why is musical timbre so hard to understand? In: S. Nielzenand & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43—53). Amsterdam: Elsevier.
- [Lakatos 2000] Lakatos, S. 2000. A common perceptual space for harmonic and percussive timbres. *Perception & Psychophysics*. 62 (7), 1426—1439, 2000.
- [Livshin & Rodet 2004] Livshin, A., & Rodet, X. 2004. Musical Instrument Identification in Continuous Recordings. In *Proceedings of the 7th Conference on Digital Audio Effects (DAFx'04), Naples, Italy, October 5 - 8, 2004*, 222—227.

# References

- [Martin & Kim 1998] Martin, K.D., & Kim, Y.E., 1998. Musical instrument identification: A pattern-recognition approach. In *Proceedings of the 136th meeting of the Acoustical Society of America*.
- [McAdams et al. 1995] McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. 1995. Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychol Res*, 58, 177—192.
- [McAdams et al. 2006] McAdams, S., Giordano, B., Susini, P., Peeters, G. 2006. A meta-analysis of acoustic correlates of timbre dimensions. *The 9<sup>th</sup> DAFx keynote slides* (<http://www.dafx.ca/slides/keynote3.pdf>).
- [Rioux et al. 2006] Rioux, V., McAdams, S., Susini, P., Peeters, G. 2006. Psycho-acoustic Timbre Descriptors, IRCAM WP2.1.5. ([http://recherche.ircam.fr/projects/cuidado/wg/management/deliverables/IRCAM\\_WP215\\_M2.pdf](http://recherche.ircam.fr/projects/cuidado/wg/management/deliverables/IRCAM_WP215_M2.pdf))
- [Saldanha & Corso 1964] Saldanha, E.L., & Corso, J.F. 1964. Timbre cues and the identification of musical instruments. *The Journal of the Acoustical Society of America*, 36(10), 2021 – 2026.
- [Wieczorkowska & Kolczynska 2008] Wieczorkowska, A., & Kolczynska, E. 2008. Quality of Musical Instrument Sound Identification for Various Levels of Accompanying Sounds. In *Lecture Notes in Computer Science*. Springer-Verlag Berlin Heidelberg, 2008, 93—103.