

# Akustische und musikalische Merkmale populärer Musik der letzten 50 Jahre

Michael Oehler<sup>1</sup>, Christoph Reuter<sup>2</sup>, Isabella Czedik-Eysenberg<sup>2</sup>, Michael Ziethen<sup>3</sup>

<sup>1</sup> Institute for Sound and Vibration Engineering, University of Applied Sciences Düsseldorf

<sup>2</sup> Musicological Department, University of Vienna

<sup>3</sup> MHMK - Macromedia University for Media and Communication, Cologne

## Hintergrund

Schellenberg und Schewe (2012) untersuchten in einer aktuellen Studie musikalische Merkmale in amerikanischer populärer Musik, die häufig mit der Vermittlung oder der Induktion von Emotionen in Verbindung gebracht werden, z.B. Tempo und Tonalität (Hunter & Schellenberg, 2010). Die Grundlage der Stichprobe waren die Jahresendcharts des *Billboard Magazine* von 1965 bis 2009. Die Autoren stellten dabei über die Jahre hinweg eine kontinuierliche Zunahme von Stücken in Moll sowie eine Reduzierung des durchschnittlichen Tempos fest. Daraus schlossen sie, dass Musik in diesem Zeitraum zunehmend trauriger klingend sowie emotional mehrdeutiger geworden sei.

## Ziele

Das Ziel unserer Studie war es zu untersuchen, ob die in der Originalstudie gefundenen Effekte auch auf eine andere Stichprobe populärer Musik, die deutschen Jahresendcharts, übertragbar sind. Darüber hinaus sollte geprüft werden, ob (psycho-)akustische Merkmale, wie *Dynamikumfang*, *Spectral Centroid*, *Rauhigkeit* oder *Inharmonizität*, deren Zusammenhang mit ausgelösten Emotionen z.T. in verschiedenen Studien nachgewiesen werden konnte (z.B. Coutinho & Dibben, 2013; Nagel et al., 2008), bessere Prädiktoren für das Aufnahmejahr sind als Tempo und Tonalität.

## Ergebnisse

### a) Replikation

Es konnte wie bei Schellenberg und Schewe (2012) eine positive Korrelation zwischen den Parametern *Jahr der Aufnahme* und *Dur-Moll-Tonalität* festgestellt werden ( $r_{pb}=.294$ ,  $p<.0001$ ), jedoch kein signifikanter Zusammenhang zwischen dem *Jahr der Aufnahme* und dem *Tempo* ( $r=.05$ ,  $p=.096$ ).

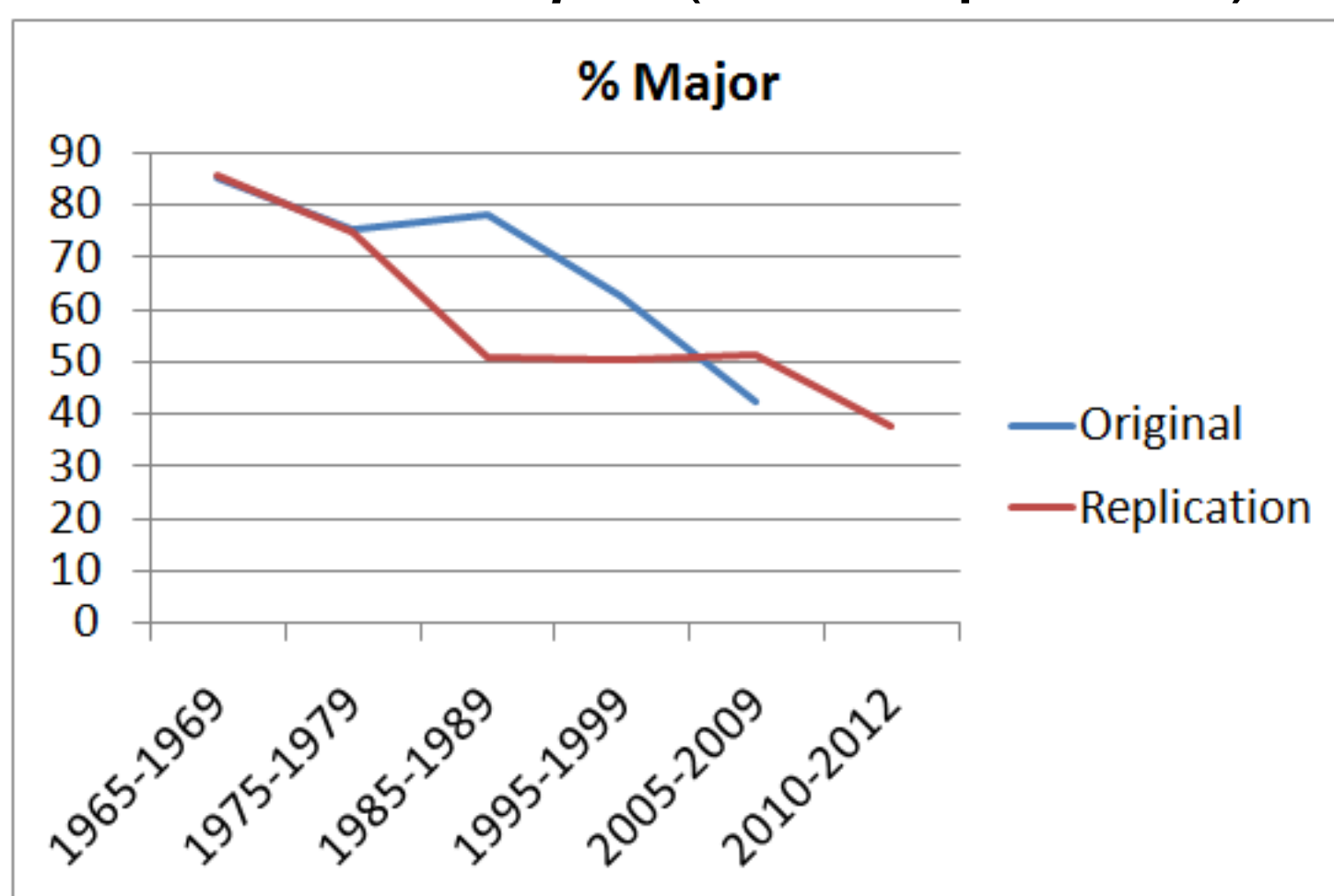


Fig 1: Experts' ratings: ratio songs in major (in %).

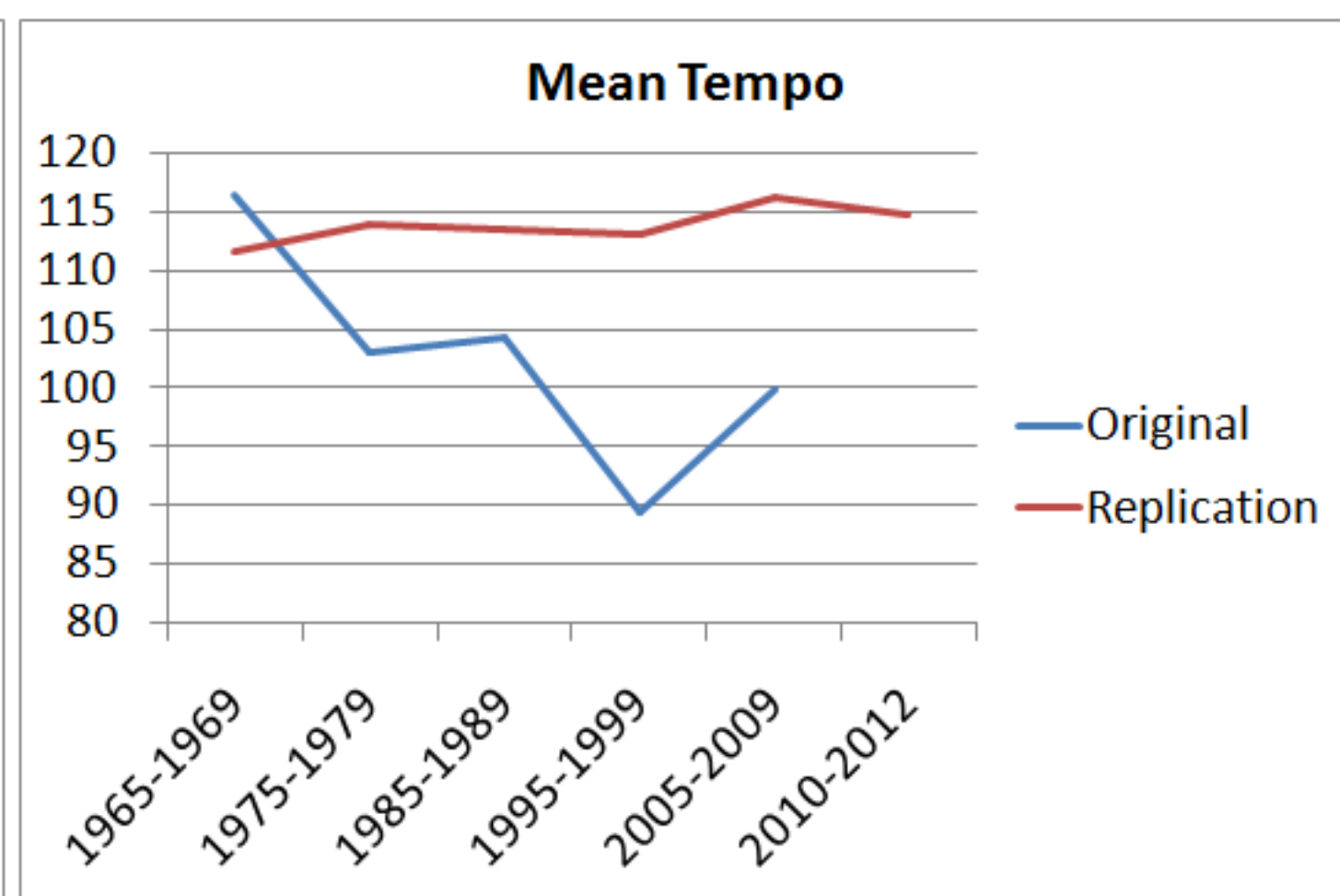


Fig 2: Experts' ratings: mean tempo (in bpm).

	Original		Replication	
	p	Beta	p	Beta
Duration	< .0001	.095	.0001	.180
Mode	< .0001	.070	.0001	.247
Tempo	< .0001	.025	.143	.041
Gender	< .0001	.011	.0001	.126

Tab 1: Multiple linear regression (org. vs. repl.).

	Beta	T	p	CI (95%)
Mode	.252	8,884,000	5,968	9,351
Duration	.176	6,237,000	.043	.082
Gender	.126	4,473,000	2,197	5,630

Tab 2: Stepwise linear regression (replication).

## Methode

Als Stichprobe wurden die deutschen von MediaControl bzw. Musikmarkt ermittelten Jahresendcharts der Jahre 1965 bis 2012 verwendet (insgesamt 1120 Stücke, d.h. jeweils die zweite Hälfte jedes Jahrzehnts, also 1965-1969, 1975-1979 etc.).

Wie in der Originalstudie erfolgte die Analyse von Tempo und Tonalität durch drei musikalische Experten. Die Interrater-Reliabilität bzw. Krippendorff's *alpha* war für den Parameter Tempo  $\alpha = .89$  sowie für dem Parameter Dur-Moll-Tonalität  $\alpha = .82$ .

Für die Extraktion der psychoakustischen Parameter kam die MIR toolbox (Lartillot & Toivainen, 2007) sowie die EchoNest-Datenbank (Andersen, 2014) zum Einsatz. Neben Einzelparametern (z.B. *Spectral Flux*) lassen sich mit Hilfe der EchoNest-Datenbank zudem komplexe Parameter bestimmen, die auf einer Kombination verschiedener psychoakustischer Merkmale beruhen (z.B. *Danceability* als Kombination aus Lautheit, Dynamikumfang, MFCC, allg. Entropie etc.).

### b) Low level features – MIR Toolbox

Die Parameter *Dynamic Range* ( $\beta = -.618$ ;  $p < .0001$ ), *Spectral Centroid* ( $\beta = .216$ ;  $p < .0001$ ), *Rauhigkeit* ( $\beta = -.221$ ;  $p < .0001$ ) sowie der Anteil von Passagen geringer Energie ( $\beta = .137$ ;  $p < .0001$ ) tragen signifikant zur Aufklärung der Varianz bei, nicht jedoch *Inharmonizität* und *Spectral Flux*.

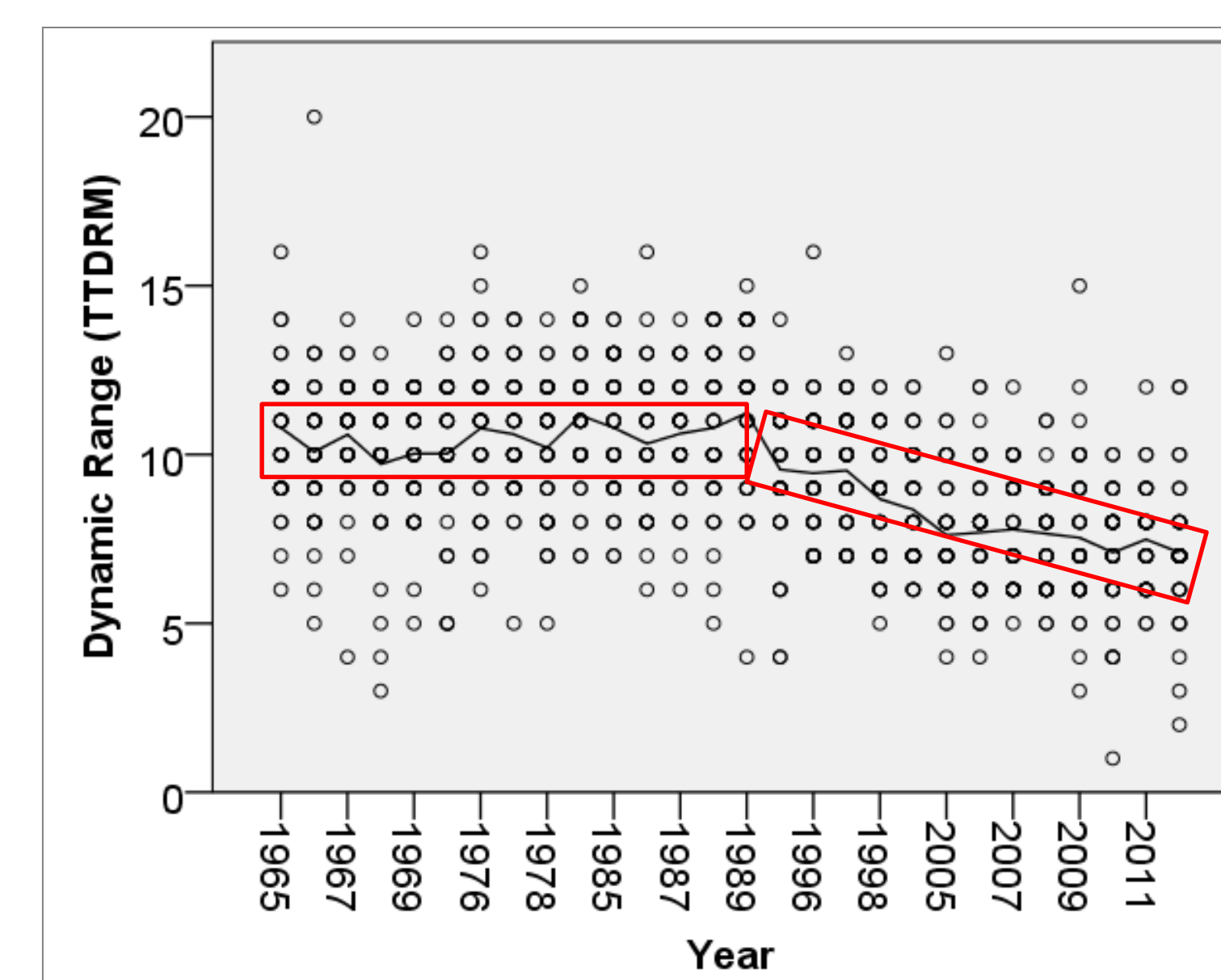


Fig 3: Dynamic Range (RMS); MIR Toolbox (decreasing since 1990;  $r = .40^{***}$ ).

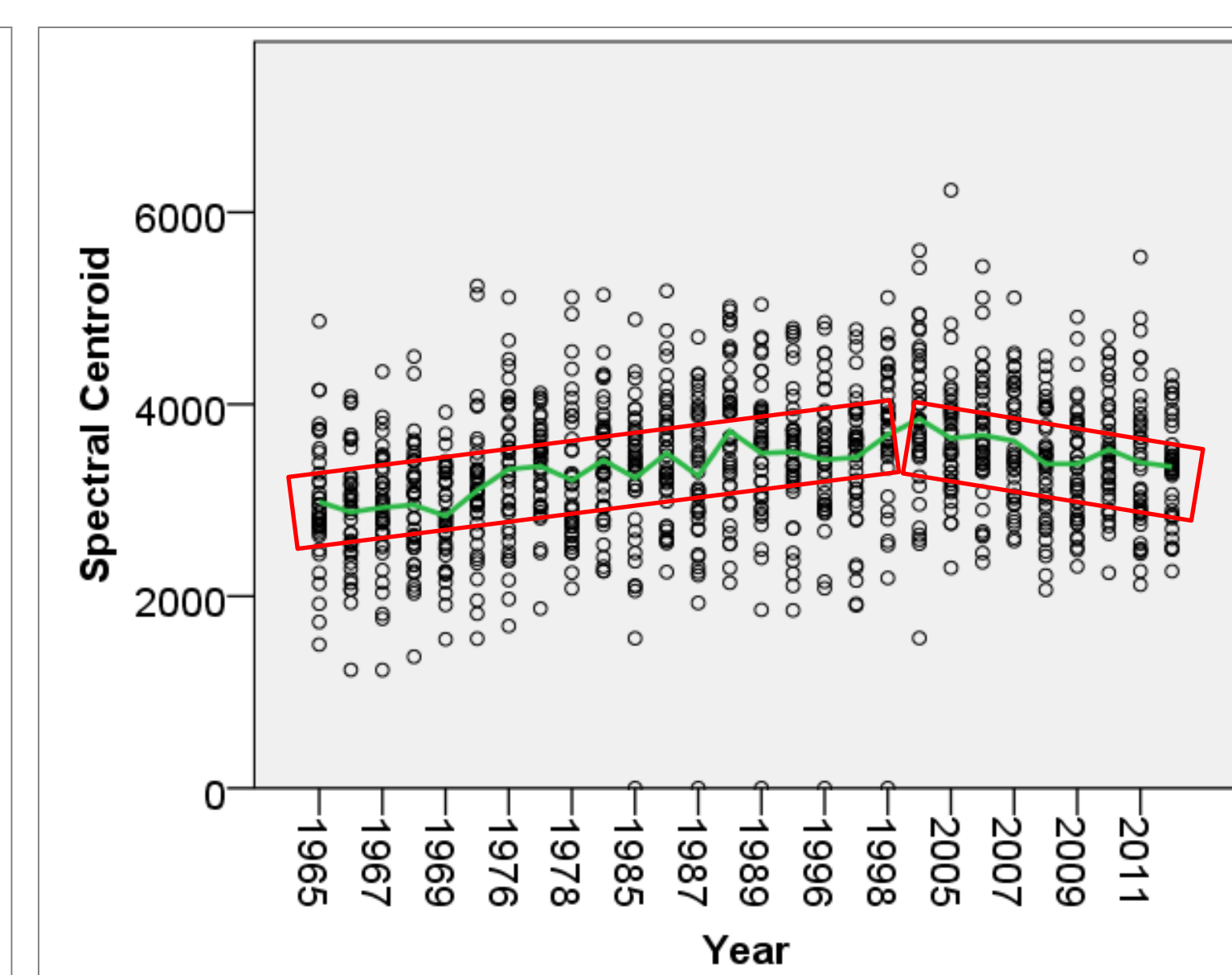


Fig 4: Spectral Centroid (in Hz); MIR Toolbox (60s – 90s:  $r = .32^{***}$ ; since 2000:  $r = -.16^{***}$ ).

	Beta	T	p	CI (95%)
Dynamic Range	-.618	-19,818	.000	-4,352 -3,568
Spectral Centroid	.216	8,506	.000	.003 .006
Roughness	-.221	-6,805	.000	-.006 -.003
Low Energy Rate	.137	5,170	.000	21,771 48,406

Tab 3: Stepwise linear regression.

## Zusammenfassung

- Die Originalstudie kann nur im Hinblick auf die kontinuierliche Zunahme von Stücken in Moll repliziert werden, nicht jedoch bezüglich der Reduzierung des durchschnittlichen Tempos seit 1965.
- Einige (psycho-)akustische Merkmale sind besonders geeignet, das Jahr der Aufnahme vorherzusagen (v.a. Spectral Centroid und Dynamic Range). Die komplexen Parameter der EchoNest Datenbank sind zwar ebenfalls geeignet, jedoch bleibt die genaue Konstruktion dieser Parameter unklar (bis dato nicht offengelegt).

## Literatur

- Andersen, J.S. (2014). Using the Echo Nest's automatically extracted music features for a musicological purpose. 4th International Workshop on Cognitive Information Processing (CIP), 26-28 May 2014, Copenhagen, Denmark.
- Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition and Emotion*, 27(4), 658-684.
- Hunter, P. G., & Schellenberg, E. G. (2011). Music and emotion. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 129-164). New York, NY: Springer.
- Lartillot, O., & Toivainen, P. (2007). A Matlab Toolbox for Musical Feature Extraction from Audio. Proceedings of the 10th International Conference on Digital Audio Effects (DAFx'07). Bordeaux, France, September 10-15, 2007.
- Nagel, F., Kopiez, R., Grewe, O., & Altenmüller, E. (2008). Psychoacoustical correlates of musically induced chills. *Musicae Scientiae*, 12(1), 101-113.
- Schellenberg, E. G., & von Scheve, C. (2012). Emotional Cues in American Popular Music: Five Decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts*, 6(3), 196-203.

### c) Low level features – EchoNest Datenbank

Auch die Parameter *Danceability* ( $\beta = .474$ ;  $p < .0001$ ) und *Valence* ( $\beta = -.482$ ;  $p < .0001$ ), tragen signifikant dazu bei, das Jahr vorherzusagen, in dem die Stücke in den Charts vertreten waren.

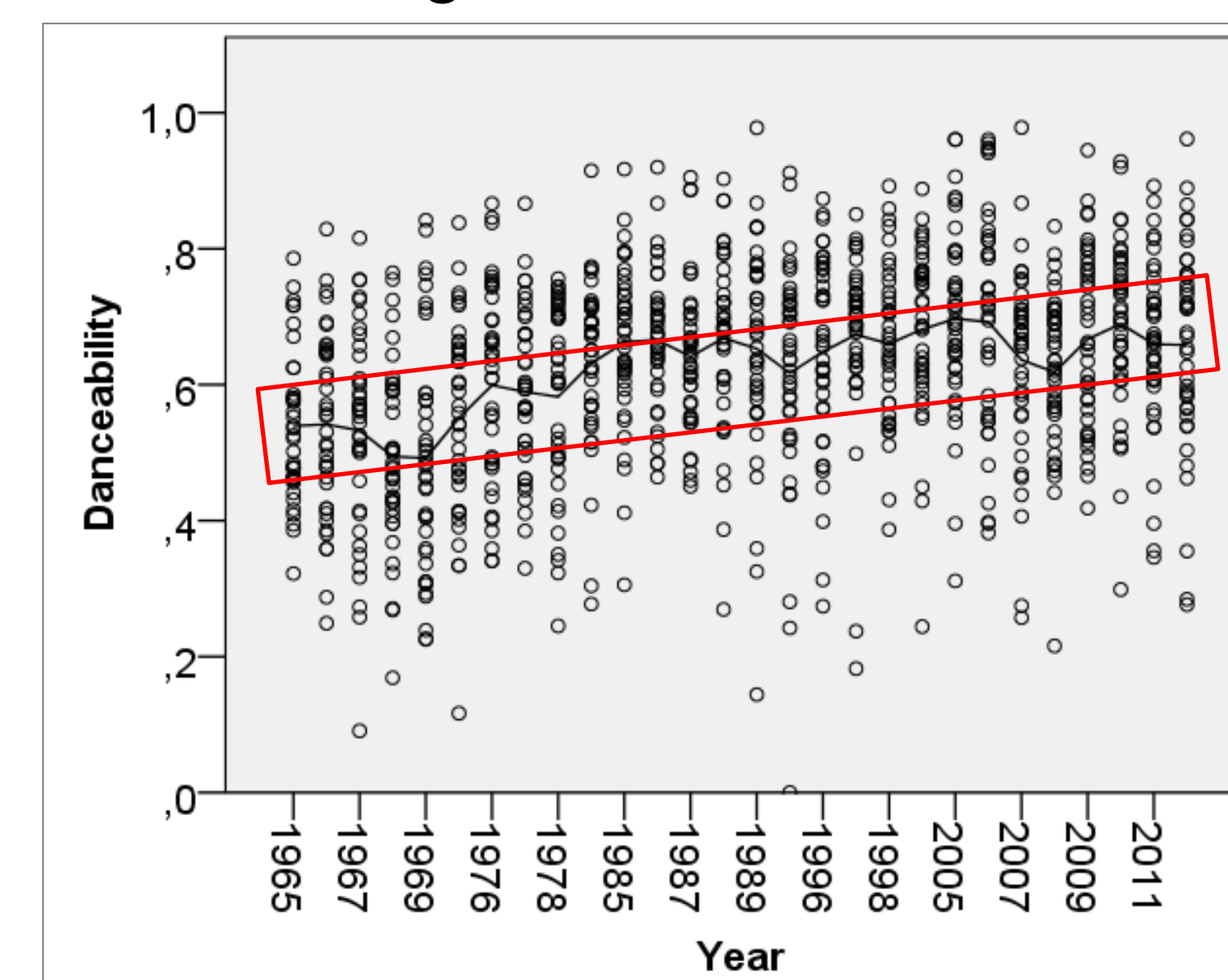


Fig 5: Danceability; EchoNest ( $r = .33^{***}$ ).

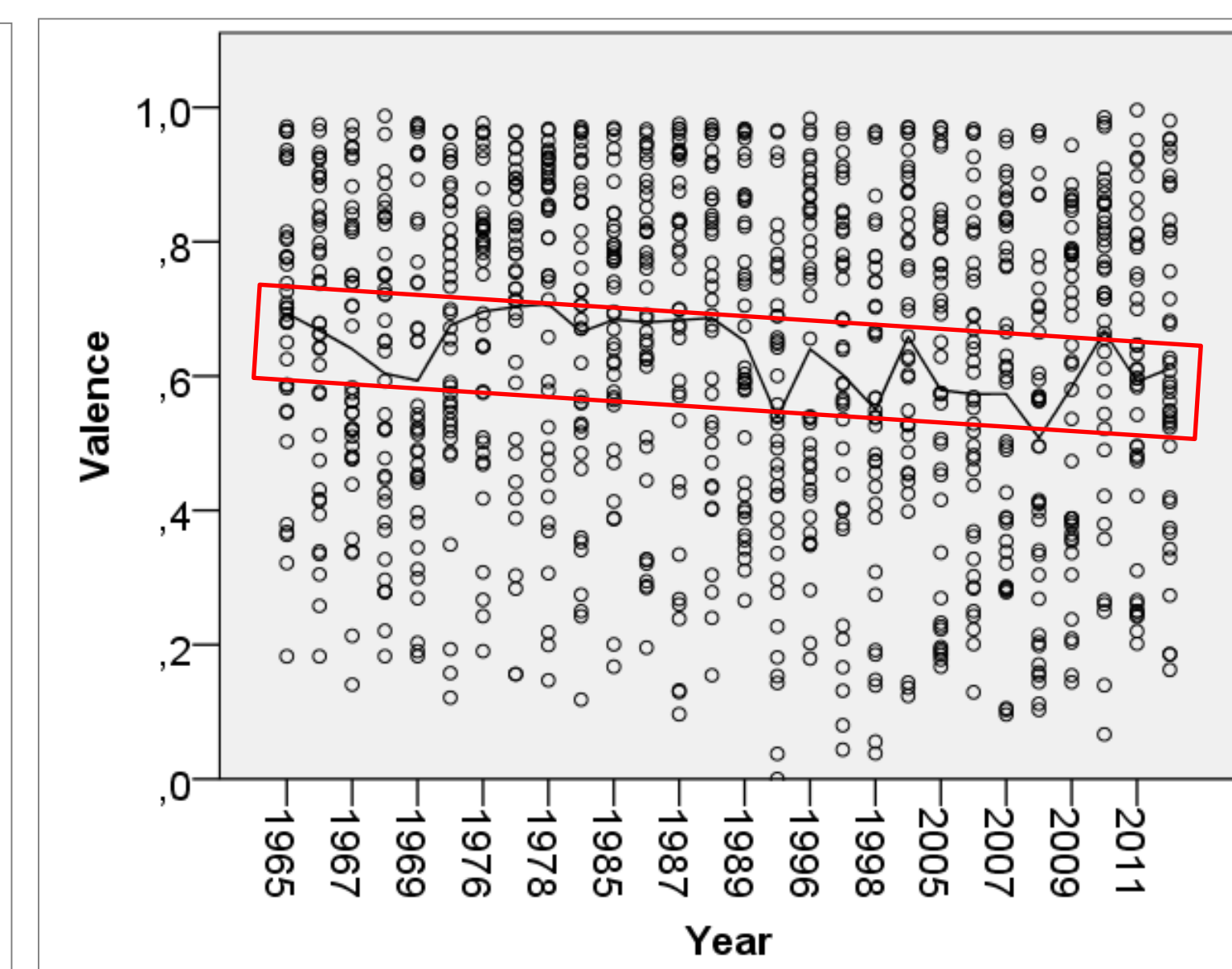


Fig 6: Valence; MIRToolbox ; EchoNest ( $r = .25^{***}$ ).