Consonance of Non-Harmonic Complex Tones: Testing the Limits of the Theory of Beats

Björn Jacobsson & Jesper Jerkert‡*

SAMMANFATTNING

Vi undersökte hur väl den enligt svävningsteorin förväntade konsonansen för icke-harmoniska komplexa toner överensstämde med ett genomfört lyssningsexperiment. Försökspersoner med olika musikalisk bakgrund ombads bedöma konsonansen hos tjugo stimuli som bestod av två sådana toner som ljöd samtidigt. Vi fann att försökspersoner som var musikakustiker bedömde stimuli i enlighet med teorin, medan andra med olika musikalisk bakgrund inte gjorde det. De senare visade däremot en tydlig korrelation mellan låg bedömd konsonans och frekvensförhållandets storlek för tonerna i stimuli, vilket kanske förklaras av de höga deltoner som fanns i dessa stimuli.

ABSTRACT

The consonance predicted by the theory of beats for non-harmonic complex tones was tested by letting test persons with different musical background evaluate the consonance of twenty stimuli consisting of two such tones sounding simultaneously. We found that the test persons who were music acousticians rated the stimuli in accordance with the theory, whereas other persons, with varied musical background, did not. However, the latter showed a significant correlation between low rated consonance and large frequency ratio for the two tones of the stimuli, possibly explained by the presence of high pitch partials.

INTRODUCTION

The theory of beats is presently widely accepted as a tool for predicting consonance and dissonance of complex tones. However, the possible applicability of this theory for non-harmonic complex tones has as yet not been fully tested. Hence, the aim of this investigation was to test the theory against experimental results using non-harmonic complex tones.

BACKGROUND

After the 1863 publication of *On the Sensations of Tone* by the German physicist-physiologist von Helmholtz, his theory of beating partials was for a long time recognized as a possible explanation of the consonance/dissonance dichotomy.[1] In the 1960's a series of experiments by Plomp & Levelt and Kameoka & Kuriyagawa seemed to show that a slight modification of that theory, namely relating the maximum dissonance of two sine tones to the critical bandwidth of the ear, was the final explanation of consonance and dissonance.[2] Terhardt has pointed out that the theory of beats indeed could explain the consonance (or dissonance) between complex tones but, as Terhardt states, it does not explain why there are such things as chord fundamentals which are said to represent the whole chord, or why tones an octave, a fifth, or a fourth apart exhibit such an extraordinary affinity, particularly in vocal music.[3] He solves these problems by adapting a two-component concept of musical consonance. However, as noted by one of the present authors, not even Terhardt's expanded consonance concept includes all the connotations attributed to the term "consonance" by musicians and musicologists.[4] One might conclude that consonance and dissonance evade exact definitions. Whether this be the case or not, it is still desirable to test the limits of the theory presently held to be the best, i.e. the theory of beats.

The theory of beats states that the dissonance of two (or more) complex tones is the sum of the beating of all adjacent partials. As Helmholtz showed, this makes perfect sense when

* E-mail: f97-bja@f.kth.se.
‡ E-mail: f97-jje@f.kth.se.
[1] Helmholtz (1877).

[2] Plomp & Levelt (1965), Kameoka & Kuriyagawa (1969a,b).
[3] Terhardt (1984).
[4] Jerkert (1999).

applied to complex tones with harmonic partials.[5] To our knowledge no listening experiments have been carried out to test the applicability of the theory for complex tones with *non*-harmonic partials. One possible reason for this could be that the theory of beats makes such convincing predictions for the consonance of harmonic tones that nobody has bothered about the non-harmonic ones.[6] Still, if this theory really catches the essence of consonance and dissonance, certainly it would be further corroborated by testing non-harmonic stimuli.

From the work of Plomp and Levelt, it is known that maximum dissonance of two beating sine tones is perceived when the frequency difference equals one fourth of the critical bandwidth of the ear.[7] The critical bandwidth is dependent on frequency in a non-linear fashion, so the modelling of a mathematical dissonance calculation is not thoroughly trivial. Recently, however, one such model has been presented by the American electro-acoustician William Sethares to be used with the software Matlab®.[8] In our investigation, the Sethares model was used to calculate the theoretical dissonance values against which to judge the empirical outcome.

METHOD

What we tried to establish was whether there is a bijection between the two sets of theoretically predicted dissonance and rated consonance. As a simplification, the bijection was assumed to be a linear function, i.e. a straight line with non-zero slope. This bijection was assumed to exist with 95% confidence if a straight line could be fitted to the data points with a slope that was non-zero with 95% confidence.

A listening experiment was performed in order to collect experimental data. The test was carried out as follows.

Subjects

Altogether 18 persons were tested. The subjects were divided into three groups reflecting their musical background. Group 1 (five subjects) consisted of people working professionally with music at the music acoustics group of the Dept. of Speech, Music and Hearing (TMH), Royal Institute of Technology (KTH), Stockholm. Group 2 (five subjects) was made up of persons well experienced in playing an instrument or singing, though only as amateurs. In the third group (eight subjects), subjects with little or no musical training were gathered. The subjects of groups 2 and 3 were mainly students.

Stimuli

Each stimulus consisted of two complex tones sounding together for five seconds. The two tones had the same spectrum, differing only with respect to frequencies. Five different inharmonic spectra were composed using Matlab®, see Table 1 for details.[9]

For each spectrum four stimuli were composed by varying the frequency ratio of the two tones involved, see Table 2.

Procedure

All subjects were tested individually. Placed in a sound proof, anechoic room, the subject was asked to rate the consonance of twenty different stimuli in two sessions, resulting in 36 sessions. The subjects were not told that the sessions in fact contained identical stimuli. All subjects were told that "consonant" should be regarded as synonymous with "euphonious", which is in concordance with the findings of van de Geer, Levelt, and Plomp.[10]

The Spruce® software package was used during the experiment, making the data collection wholly automatized. The subjects chose

---

[5] See the famous figure in Helmholtz (1877), p. 193.

[6] However, no extensive literature search has been made to confirm that no experiments have been carried out. [*Addendum*: Actually, experiments very similar to ours have been reported once. A team of British researchers asked subjects to rate sounds extracted from a modified form of Chladni's law, resembling bells, gongs and cymbals, in terms of pleasantness and interestingness. They found a weak correlation or no correlation between pleasantness and calculated dissonance, so their results are in concordance with ours. See G. M. Swallowe et al. (1997): "On consonance: Pleasantness and interestingness of four component complex tones," *Acustica* **83**, 897-902.]

[7] Plomp & Levelt (1965).

[8] Sethares (1993) and (1997). We have not found it necessary to examine very carefully Sethares' mathematical model, but those who want to should be warned that his mathematical discussion in the 1993 paper is somewhat obscure.

[9] Tables are found in appendix A.

[10] van de Geer, Levelt & Plomp (1962).

which stimulus to play by clicking one of twenty icons randomly placed in a field on the screen. Rating was executed by dragging and dropping the icon into another field with a scale where one end represented maximum consonance and the other minimum consonance. It was possible to place several stimuli beside each other, representing the same consonance or dissonance. The stimuli could be replayed at any time, allowing the rating to be adjusted until the subject was satisfied. Finally the subjects were asked to give some personal opinions on the test.

CALCULATIONS AND RESULTS

The software which was used for the tests collected the data into a file, rating the consonance on a scale from 0 to 1000, where 1000 represented highest consonance. Using Matlab® the data were edited and analyzed in the following ways:

First the two sessions of each subject were compared, measuring the discrepancy of the ratings of the same stimuli in the two sessions. Those subjects who had a mean discrepancy of over 200 were excluded. One subject, belonging to group 1, was excluded in this way.[11]

Secondly, the data were compared to the results predicted by the theory of beats. The theoretical values were calculated using Sethares' model.[8] For the data points (x_i^j, y_i^j), x_i^j represented rated consonance, y_i^j theoretical dissonance, i = 1, …, 20 were the number of stimuli and j = 1, …, 34 the number of sessions. For fixed j, the y_i^j's were assumed to be observations of the normal random variable Y_i^j, where $Y_i^j \in N(m_i^j, \sigma^j)$, $m_i^j = \alpha^j + \beta^j x_i^j$, allowing linear regression to be used to fit a curve to the data points for each j. The lines obtained are plotted in Figure 1, both all together and separated in the three groups of subjects.[12]

The β^j's are thus observations of a normal random variable, $N(\beta^j, \frac{\sigma^j}{\sqrt{\sum_i (x_i^j - \bar{x}^j)^2}})$.[13]

Not knowing if the results from the sessions represented the same theoretical slope β, each β^j was analyzed separately.

To test the hypothesis H_0: $\beta = 0$, a 95% confidence interval for each β^j was calculated as presented in Figure 2.[13] H_0 can be rejected with 95% confidence if the β^j for a fixed j is outside the interval.[14] Thus the hypothesis was rejected for 5 of 8 sessions in group 1, 0 of 10 sessions in group 2, and 1 of 16 sessions in group 3.

When commenting on the test, many of the subjects expressed opinions, which can be summarized as finding stimuli containing partials with high pitches more annoying than others. Hence, the collected data were plotted versus the frequency ratio of the two tones of the stimulus concerned. The data points were analyzed in the same manner as above, resulting in Figures 3 and 4. Thus the hypothesis that the slope was zero could be rejected for 2 of 8 sessions in group 1, 7 of 10 sessions in group 2, and 11 of 16 sessions in group 3.

As will be discussed below, we found a correlation between rated consonance and frequency ratio of the tones. To examine a possible correlation between theoretically predicted dissonance and frequency ratio, the two factors were plotted versus each other for each stimulus, see Figure 5.

DISCUSSION

For a majority of the music acousticians theoretical dissonance and rated consonance are correlated, whereas no correlation could be found between theoretical dissonance and frequency ratio. For most subjects in groups 2 and 3 the result was the opposite, no correlation could be found between theoretical dissonance and rated consonance, while theoretical dissonance and frequency ratio were, in fact, correlated.

---

[11] This was not surprising, as the subject in question stated that s/he rated the stimuli very differently in the two sessions.

[12] Figures are found in appendix B.

[13] Blom (1989), p. 285.

[14] Blom (1989), p. 260.

As shown by Figure 5 the distribution of intervals of the two tones and the theoretical dissonances does not show any correlation. This relation should therefore not interfere with the results above.

In listening experiments it is customary to include some trial stimuli before the real test begins. The automatized procedure used in this study, allowing the subject to play each stimulus an unlimited number of times, made such arrangements unnecessary.

Why do the subjects of groups 2 and 3 dislike the stimuli containing large frequency ratios? Possibly, the mere presence of high pitch partials makes the stimuli shrill and ill-sounding. However, such a hypothesis needs further testing to be accepted or rejected.

Our investigation comprises the analysis of ratings from 17 subjects, divided into three groups. While this gives very few subjects in each group, the relative consistency of the data within the groups makes us believe that our conclusions are still well-founded.

An interesting observation is that subjects who were familiar with the theory of beats rated the stimuli accordingly, while those unacquainted with the theory did not. This raises the question whether this theory has any relevance at all, since then it would constitute a *pre*scription rather than a *de*scription of human perception.

CONCLUSIONS

According to our investigation the theory of beats does not explain the concept of consonance entirely. When listening to non-harmonic complex tones, other factors appear to be important in judging the consonance. One such factor seems to be the presence of high pitch partials, though this has to be further investigated. In any case, the incongruity of our experimental results with the theory of beats highlights the need for a more comprehensive theory of consonance and dissonance.

ACKNOWLEDGMENTS

This work was carried out as a student project in the undergraduate course Music Acoustics at the Dept. of Speech, Music and Hearing (TMH), Royal Institute of Technology (KTH), Stockholm. Primarily, we would like to thank our supervisor Anders Askenfelt for valuable criticism. Our thanks also go to the Department for letting us make use of the fine equipment of the acoustics laboratory. Moreover, we would like to express our gratitude to Svante Granqvist, whose excellent Spruce® software substantially facilitated the collection of data. Last but not least we thank all those people who spent some twenty minutes each judging the sometimes ill-sounding stimuli; your contributions were essential for the investigation.

REFERENCES

Blom, G. (1989): *Sannolikhetsteori och statistikteori med tillämpningar*. Lund: Studentlitteratur, 1989.

van de Geer, J. P.; Levelt, W. J. M. & Plomp, R. (1962): "The connotation of musical consonance," *Acta Psychol.* **20** (1962), 308-319.

Helmholtz, H. von (1877): *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. New York: Dover, 1954. (Original German ed.: *Die Lehre von Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Braunschweig: Vieweg, 1/1863, 4/1877.)

Jerkert, J. (1999): *Vad är konsonans och dissonans? Musikvetenskapliga och naturvetenskapliga perspektiv* ["What is Consonance and Dissonance? Musicological and Scientific Perspectives"]. Unpublished B.A. thesis (in Swedish), Dept. of Musicology, Univ. of Stockholm, 1999.

Kameoka, A. & Kuriyagawa, M. (1969a): "Consonance theory part I: Consonance of dyads," *J. Acoust. Soc. Am.* **45**(6) (1969), 1451-1459.

——— (1969b): "Consonance theory part II: Consonance of complex tones and its calculation method," *J. Acoust. Soc. Am.* **45**(6) (1969), 1460-1469.

Plomp, R. & Levelt, W. J. M. (1965): "Tonal consonance and critical bandwidth," *J. Acoust. Soc. Am.* **38**(4) (1965), 548-560.

Sethares, W. A. (1993): "Local consonance and the relationship between timbre and scale," *J. Acoust. Soc. Am.* **94**(3) (1993), 1218-1228.

——— (1997): *Tuning, Timbre, Spectrum, Scale*. London: Springer, 1997.

Terhardt, E. (1984): "The concept of musical consonance: A link between music and psychoacoustics," *Music Perception* **1**(3) (1984), 276-295.

APPENDIX A. TABLES

Name of spectrum	*Frequencies of partials of the lower tone* $f_0 = 256\ \mathrm{Hz}$, n = number of partial	*Relative amplitude of partials*
A	$f_0 (2.1)^{\ln n / \ln 2}$, $1 \le n \le 11$	$(0.7)^{n-1}$, $1 \le n \le 11$
B	$f_0 (1.85)^{\ln n / \ln 2}$, $1 \le n \le 11$	$(0.8)^{n-1}$, $1 \le n \le 11$
C	$f_0 (1, 2.40, 3.83, 5.52, 7.02, 8.65)$ or, exactly, $f_0 (1, j_{01}, j_{11}, j_{02}, j_{12}, j_{03})$	$(0.8)^{n-1}$, $1 \le n \le 6$
D	$f_0 (1, 1.73, 3.26, 4.11, 4.58, 7.39, 8.20)$	$(0.8)^{n-1}$, $1 \le n \le 7$
E	$f_0 (1, 1.44, 2.72, 3.14, 4.82, 7.37)$ or, exactly, $f_0 (1, \pi^{1/\pi}, e, \pi, e^e / \pi, e^\pi / \pi)$	$(0.8)^{n-1}$, $1 \le n \le 6$

Table 1. Composition of spectra.

A and B represent stretched and compressed spectra, respectively; in both frequency formulas, substituting 2.1 and 1.85 with 2 will yield harmonic spectra. The frequency ratios of spectrum C represented zeros of the Bessel functions of the first kind of order 0 and 1 in the interval [0,9]. Spectrum D was composed arbitrarily by the authors, and so was spectrum E, by combining the numbers e and π.

Name of stimulus	*Spectrum type of tones*	*Frequency ratio* r *of higher and lower tone*
1	A	1.362
2	A	1.542
3	A	2.001
4	A	2.100
5	B	1.291
6	B	1.433
7	B	1.756
8	B	1.850
9	C	1.271
10	C	1.517
11	C	1.831
12	C	2.225
13	D	1.125
14	D	1.405
15	D	1.593
16	D	1.831
17	E	1.175
18	E	1.440
19	E	1.576
20	E	1.888

Table 2. The composition of stimuli 1 through 20.

For all stimuli the lower tone consisted of partials with frequencies as indicated in Table 1. The frequency of the first partial of the lower tone was always 256 Hz, whereas the first partial of the higher tone had a frequency of 256· r Hz.

APPENDIX B. FIGURES

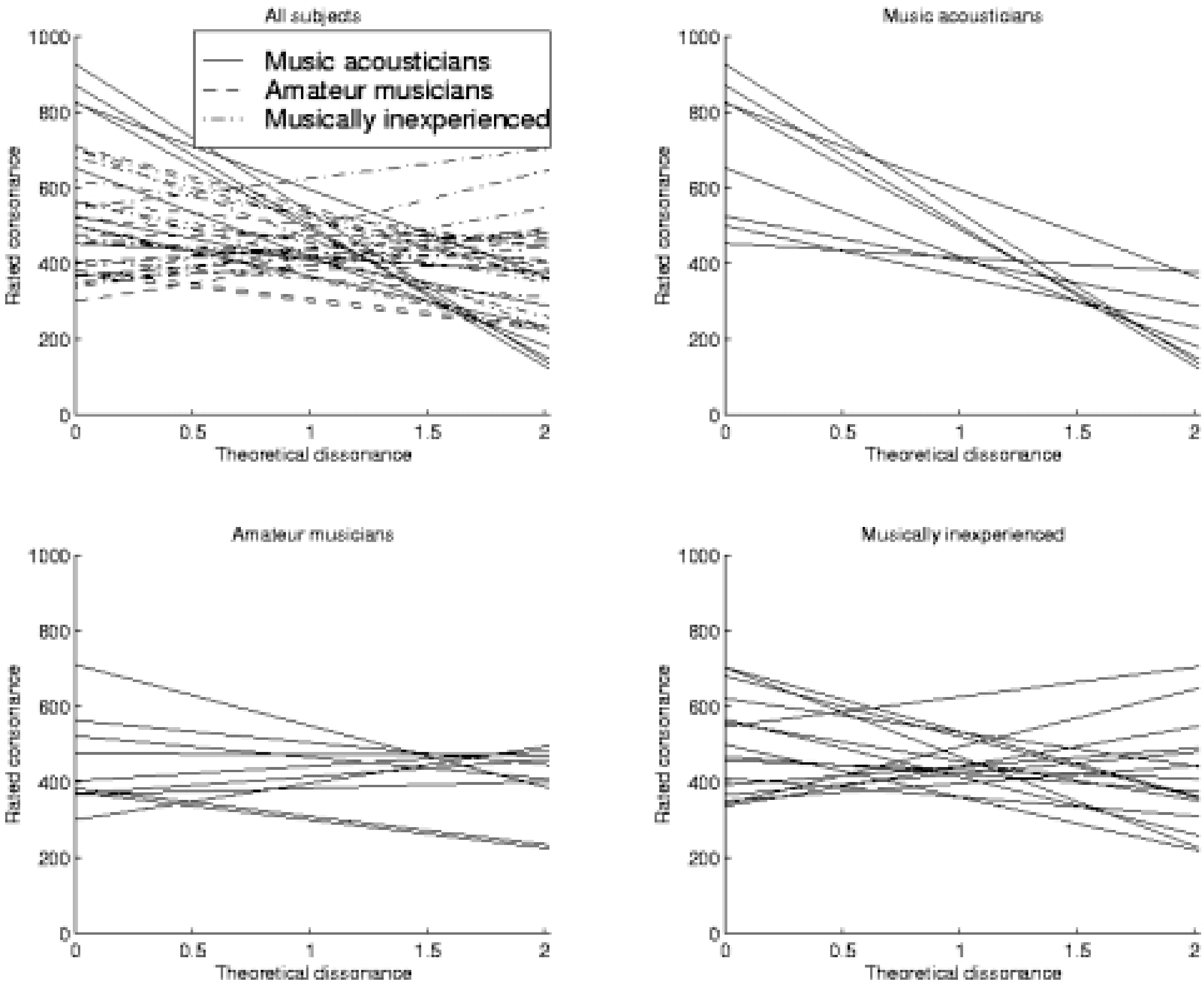

Figure 1. Rated consonances as a function of theoretically calculated dissonance.

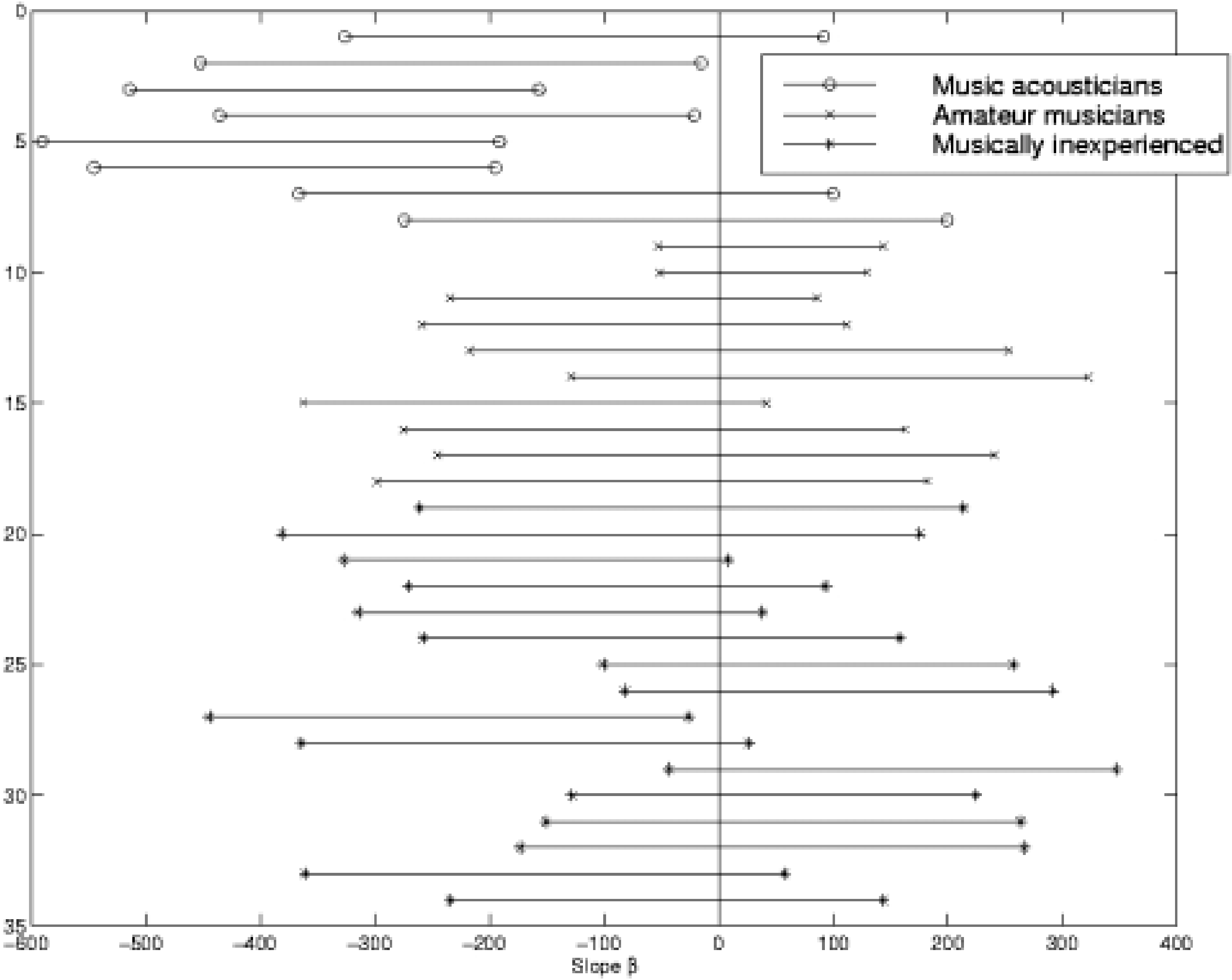

Figure 2. 95% confidence intervals of the slopes β^j for the correlation between theoretical dissonance and rated consonance for all sessions.

The numbers of sessions where the interval does not include zero are as follows: 5 of 8 for the music acousticians, 0 of 10 for the amateur musicians and 1 of 16 for the musically inexperienced.

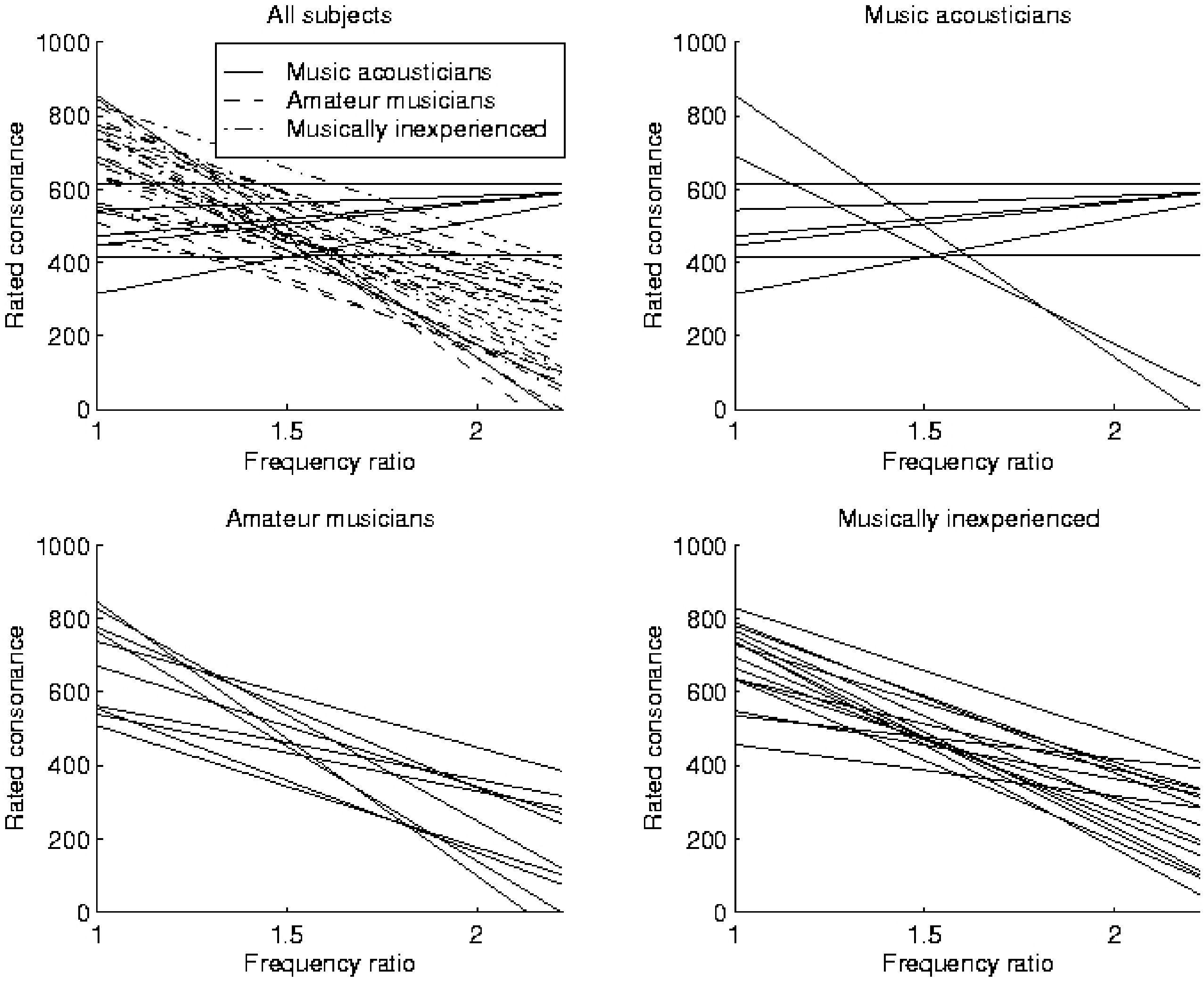

Figure 3. Rated consonances as function of frequency ratio of the two tones forming the stimulus.

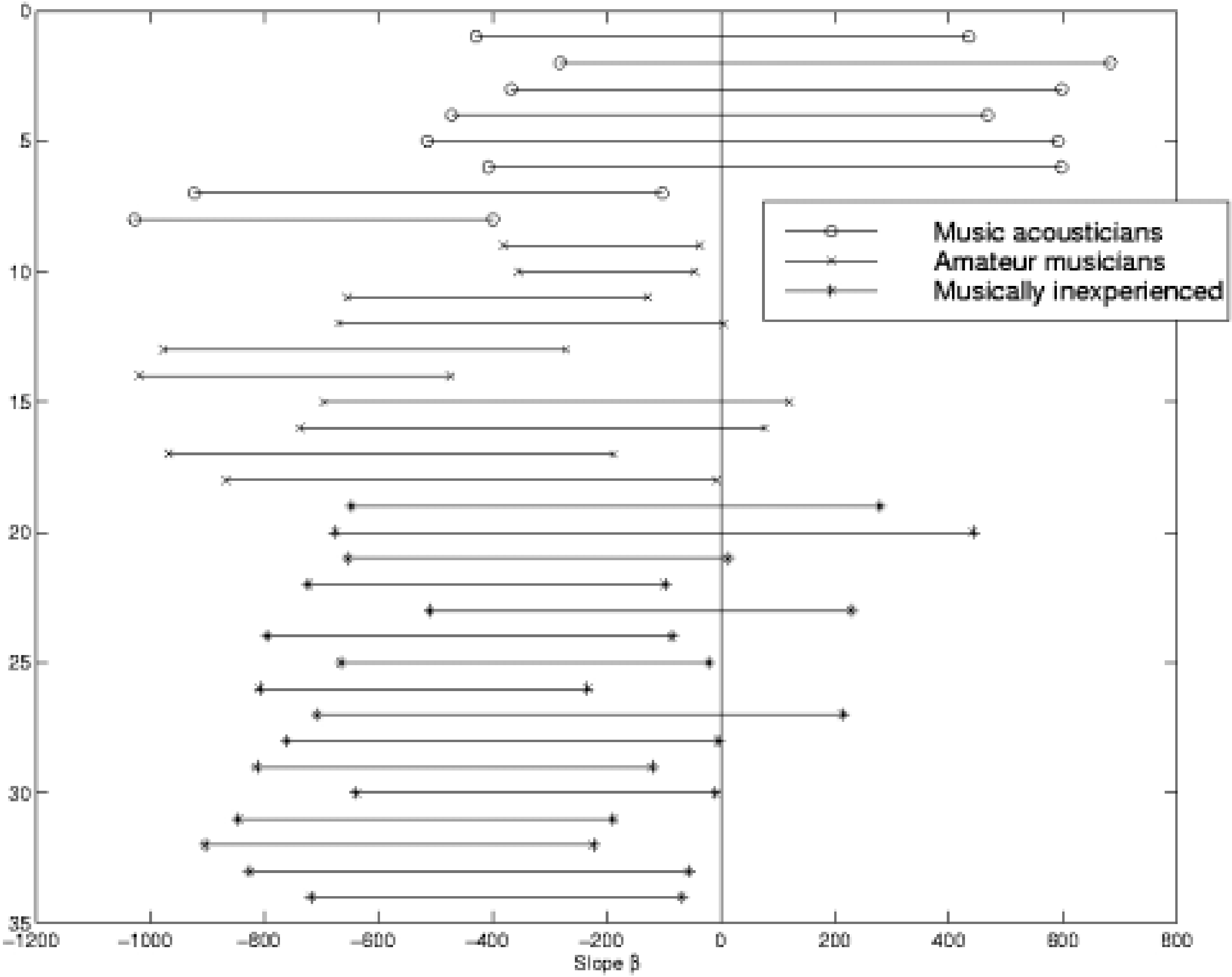

Figure 4. 95% confidence intervals of the slopes β^j for the correlation between theoretical dissonance and frequency ratio for all sessions.

The numbers of sessions where the interval does not include zero are as follows: 2 of 8 for the music acousticians, 7 of 10 for the amateur musicians and 11 of 16 for the musically inexperienced.

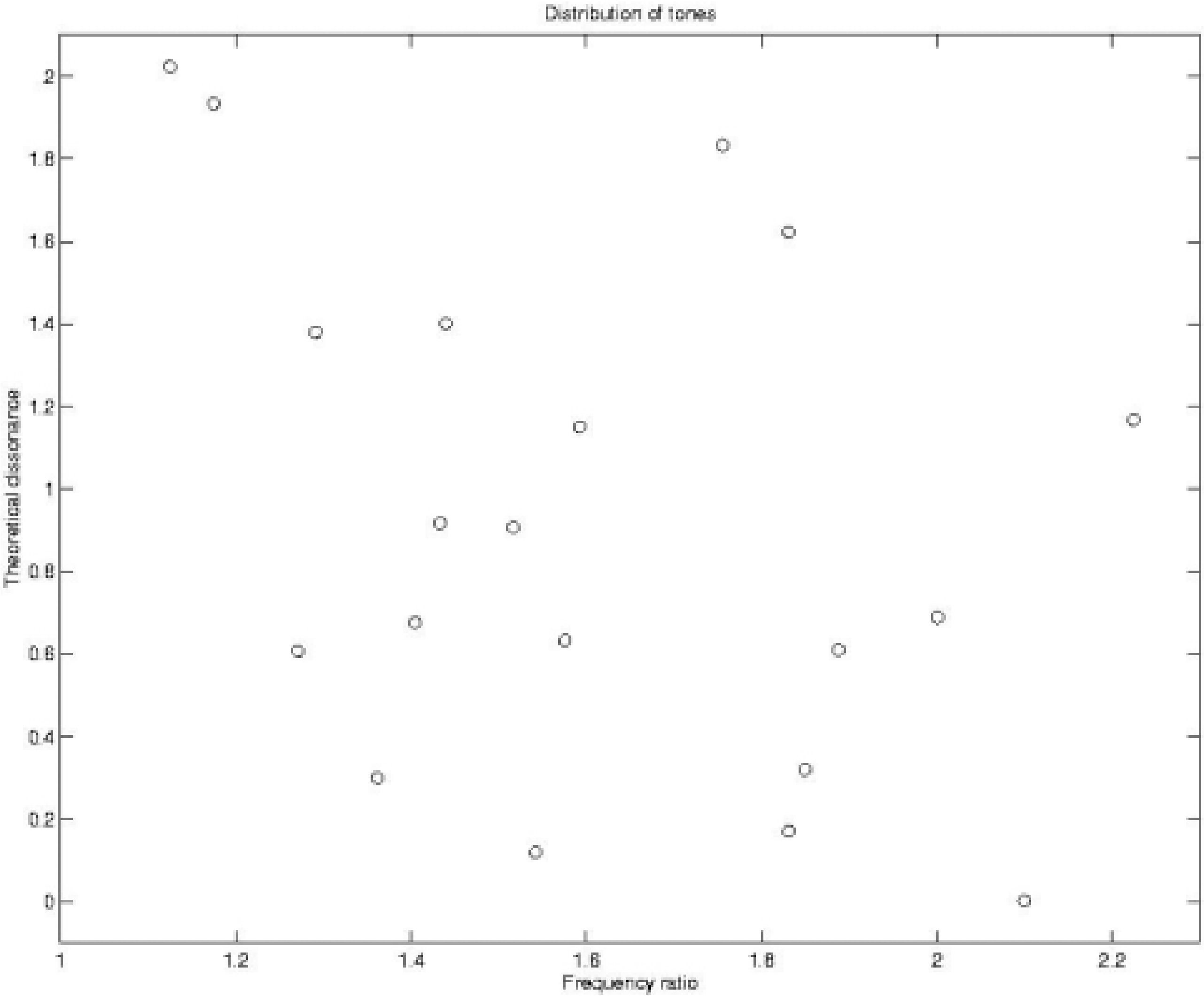

Figure 5. Distribution of tones comparing the frequency ratio of the two stimulus tones with the theoretical dissonance.

No tendency that stimuli with larger frequency ratios exhibit greater theoretical dissonance can be detected. A straight line fitting the points in a least square sense will have a slightly negative slope, i.e. larger frequency ratio gives less theoretical dissonance.