HARMONY PERCEPTION BY PERIODICITY DETECTION

Frieder Stolzenburg

Harz University of Applied Sciences, Automation and Computer Sciences Department, Friedrichstr. 57-59, 38855 Wernigerode, Germany, e-mail: fstolzenburg@hs-harz.de

1. INTRODUCTION

Numerous approaches tackle the question, how musical consonance/dissonance perception can be explained. Many corresponding empirical studies reveal a clear preference ordering on the perceived consonance/dissonance of common triads in Western music, e.g. major < minor (Roberts, 1986). Early mathematical models expressed musical intervals, i.e. the distance between two pitches, by simple fractions. Newer explanations base on the notion of dissonance, roughness, instability or tension (Helmholtz, 1862; Hutchinson and Knopoff, 1978, 1979; Cook and Fujisawa, 2006). They correlate better to empirical results on harmony perception, but still can be improved.

A general theory of harmony perception should be applicable to musical harmonies in a broad sense, i.e. to chords and scales. It should consider results from neuroacoustics and psychophysics on auditory processing, e.g. that periodicities of complex chords can be detected in the brain (Langner, 1997; Lee et al., 2009). Furthermore, the correlation between the predicted and the perceived consonance/dissonance from empirical studies should be the highest possible. Several empirical experiments on harmony perception have been conducted, where participants are asked to listen to and rate musical harmonies. The corresponding results (Malmberg, 1918; Johnson-Laird et al., 2012; Temperley and Tan, 2013) can be used to evaluate theories on harmony perception.

2. MAIN CONTRIBUTION

We apply recent results from psychophysics and neuroacoustics consistently, obtaining a fully computational theory of consonance/dissonance perception. We focus on periodicity detection including autocorrelation analysis, which can be identified as a fundamental mechanism to music perception, and exploit in addition that the just noticeable difference of human pitch perception is about 1% for the musically important low frequency range (Zwicker et al., 1957). We transfer the concept of periodicity pitch to chords and scales by considering relative periodicity, i.e. the approximated ratio of the period length of the chord relative to the period length of its lowest tone component. The hypothesis is that the perceived consonance of a musical harmony decreases with the relative (logarithmic) periodicity.

In this context, we adopt the equal temperament as reference system for tunings here. The frequency ratios in equal temperament are irrational numbers (except for the unison and its octaves), but for computing periodicity they must be fractions. We thus consider tunings with rational frequency ratios. The oldest tuning with this property is probably the Pythagorean tuning which, however, is not in line with the results of psychophysics, in particular, that the just noticeable difference of pitch perception is about 1%. We therefore use the rational tuning, which takes the fractions with smallest possible denominators, such that the relative deviation with respect to equal temperament is just below a given percentage d. Its frequency ratios deviate only slightly from just tuning, namely only for the tritone and the minor seventh. Just tuning can be understood as rational tuning with only slightly greater maximal relative deviation d=1.1%. We adopt it as well as the rational tuning as underlying tuning in our analyses. Table 1 shows the respective frequency ratios.

interval	k	rational tuning	just tuning
unison	0	1/1 (0,00%)	1/1 (0,00%)
minor second	1	16/15 (0.68%)	16/15 (0.68%)
major second	2	9/8 (0.23%)	9/8 (0.23%)
minor third	3	6/5 (0.91%)	6/5 (0.91%)
major third	4	5/4 (-0.79%)	5/4 (-0.79%)
perfect fourth	5	4/3 (-0.11%)	4/3 (-0.11%)
tritone	6	17/12 (0.17)	7/5 (-1.01%)
perfect fifth	7	3/2 (0.11%)	3/2 (0.11%)
minor sixth	8	8/5 (0.79%)	8/5 (0.79%)
major sixth	9	5/3 (-0.9%)	5/3 (-0.9%)
minor seventh	10	16/9 (-0.23%)	9/5 (1.02%)
major seventh	11	15/8 (-0.68%)	15/8 (-0.68%)
octave	12	2/1 (0.00%)	2/1 (0.00%)

Table 1: Table of frequency ratios for different tunings. Here, k denotes the number of the semitone corresponding to the given interval. In parentheses, the relative deviation of the respective frequency ratio from equal temperament is shown. The maximal deviation d for the rational tuning listed here is 1%.

3. RESULTS

The predictions of the periodicity-based approach obtained for dyads, common triads, and diatonic scales all show highest correlation with empirical results (Schwartz et al., 2003; Johnson-Laird et al., 2012; Temperley and Tan, 2013), not only with respect to the ranks, but also with the ordinal values of the empirical ratings of musical consonance, in particular, when logarithmic periodicity is employed, which can be motivated by the topological organization of the periodicity coding in the brain. Interestingly, the logarithmic periodicity of the complete chromatic scale is within the biological bound of 8 octaves, which can be represented in the neuronal periodicity map in the brain.

In our analyses, we correlate the empirical and the theoretical ratings of harmonies. Since in most cases only data on the ranking of harmonies is available, we mainly correlate rankings. Nevertheless, correlating concrete numerical values yields additional interesting insights (see below). For the sake of simplicity and consistency, we always compute Pearson's correlation coefficient *r*, which coincides with Spearman's rank correlation coefficient on rankings, provided that there are not too many bindings, i.e. duplicate values.

Table 2 shows the perceived and computed relative consonance of dyads (intervals). The correlations of the empirical rating with the sonance factor and with relative or logarithmic periodicity show the highest correlation (r=.982). Table 3 shows the perceived and computed relative consonance of common triads. There are several empirical studies on the perception of common triads. But since the experiments conducted by Johnson-Laird et al. (2012) are the most comprehensive, because they examined all 55 possible three-note chords, we adopt this study as reference for the empirical ranking here. Nonetheless, all studies are consistent with the following preference ordering on triads: major < minor < suspended < diminished < augmented, at least for chords in root position. However, the ordinal ratings of minor and suspended chords do not differ very much. Again, the analysis reveals highest correlations for relative and logarithmic periodicity, if the underlying tuning is psychophysically motivated. Roughness (Hutchinson and Knopoff, 1978, 1979) and the sonance factor (Hofmann-Engl, 2004, 2008) yield relatively bad predictions on the perceived consonance of common triads.

The data sets in Johnson-Laird et al. (2012) suggest further investigations. So, Table 4 shows the analysis of all possible three-tone chords in root position. As one can see, the correlation between the empirical rating with the predictions of the dual-process theory is very high (r=.916). This also holds for logarithmic periodicity but not that much for relative periodicity, in particular, if the correlation between the ordinal rating and the concrete periodicity values are taken. This justifies our preference of logarithmic to relative periodicity, because the former notion is motivated more by neuroacoustical results, namely that the spatial structure of the periodicity-pitch representation in the brain is organised as a logarithmic periodicity map.

Temperley and Tan (2013) investigate the perceived consonance of diatonic scales. Table 5 lists all classical church modes, i.e. the diatonic scale and its inversions. The cognitive model on the perception of diatonic scales introduced by Temperley and Tan (2013) results in a 100% correlation with the empirical data. Although the correlation for logarithmic periodicity obviously is not that good, it shows still high correlation. Nonetheless, the major scale (Ionian) appears in the front rank of 462 possible scales with 7 out of 12 tones with respect to relative and logarithmic periodicity. In addition, in contrast to more cognitive theories on harmony perception, the periodicity-based approach introduced in this paper does not presuppose any principles of tonal music, e.g. the existence of diatonic scales or the common use of the major triad. They can be derived from underlying, more primitive mechanisms, namely periodicity detection in the human (as well as animal) brain.

4. REFERENCES

Bigand, E., Parncutt, R., and Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Perception and Psychophysics, 58(1):125–141.

Cook, N. D. and Fujisawa, T. X. (2006). The psychophysics of harmony perception: Harmony is a three-tone phenomenon. Empirical Musicology Review, 1(2):1–21.

Gill, K. Z. and Purves, D. (2009). A biological rationale for musical scales. PLoS ONE, 4(12):e8144.

Johnson-Laird, P. N., Kang, O. E., and Leong, Y. C. (2012). On musical dissonance. Music Perception, 30(1):19–35.

Hofmann-Engl, L. J. (2004). Virtual pitch and its application to contemporary harmonic analysis. Chameleon Group Online Publication.

Hofmann-Engl, L. J. (2008). Virtual pitch and the classification of chords in minor and major keys. In Miyazaki, K., Hiraga, Y., Adachi, M., Nakajima, Y., and Tsuzaki, M., editors, Proceedings of 10th International Conference on Music Perception and Cognition, Sapporo, Japan.

Hutchinson, W. and Knopoff, L. (1978). The acoustic component of western consonance. Interface/Journal of New Music Research, 7(1):1–29.

Hutchinson, W. and Knopoff, L. (1979). The significance of the acoustic component of consonance in Western triad. Journal of Musicological Research, 3(1-2):5–22.

Langner, G. (1997). Temporal processing of pitch in the auditory system. Journal of New Music Research, 26:116–132.

Lee, K. M., Skoe, E., Kraus, N., and Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. The Journal of Neuroscience, 29(18):5832–5840.

Malmberg, C. F. (1918). The perception of consonance and dissonance. Psychological Monographs, 25(2):93–133.

Schwartz, D. A., Howe, C. Q., and Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. The Journal of Neuroscience, 23(18):7160–7168.

Temperley, D. and Tan, D. (2013). Emotional connotations of diatonic modes. Music Perception, 30(3):237–57.

Zwicker, E., Flottorp, G., and Stevens, S. S. (1957). Critical band width in loudness summation. Journal of the Acoustical Society of America, 29(5):548–557.

Further information: arxiv.org/abs/1306.6458

interval		emp. rank	roughness	sonance factor	similarity	rel. periodicity
unison	{0, 0}	1	2 (0.0019)	1-2 (1.000)	1-2 (100.00%)	1-2 (1.0)
octave	{0, 12}	2	1 (0.0014)	1-2 (1.000)	1-2 (100.00%)	1-2 (1.0)
perfect fifth	{0, 7}	3	3 (0.0221)	3 (0.737)	3 (66.67%)	3 (2.0)
perfect fourth	{0, 5}	4	4 (0.0451)	4 (0.701)	4 (50.00%)	4-5 (3.0)
major third	{0, 4}	5	6 (0.0551)	5 (0.570)	6 (40.00%)	6 (4.0)
major sixth	{0, 9}	6	5 (0.0477)	6 (0.526)	5 (46.67%)	4-5 (3.0)
minor sixth	{0, 8}	7	7 (0.0843)	7 (0.520)	9 (30.00%)	7-8 (5.0)
minor third	{0, 3}	8	10 (0.1109)	8 (0.495)	7 (33.33%)	7-8 (5.0)
tritone	{0, 6}	9	8 (0.0930)	11 (0.327)	8 (31.43%)	9 (6.0)
minor seventh	{0, 10}	10	9 (0.0998)	9 (0.449)	10 (28.89%)	10 (7.0)
major second	{0, 2}	11	12 (0.2690)	10 (0.393)	11 (22.22%)	12 (8.5)
major seventh	{0, 11}	12	11 (0.2312)	12 (0.242)	12 (18.33%)	11 (8.0)
minor second	{0, 1}	13	13 (0.4886)	13 (0.183)	13 (12.50%)	13 (15.0)
correlation *r*			.967	.982	.977	.982

Table 2: Consonance rankings of dyads. The respective numbers of semitones with respect to the Western twelve-tone system are given in braces, raw values of the respective measures in parentheses. The empirical rank is the average rank according to the summary given by Schwartz et al. (2003, Figure 6). The roughness values are taken from Hutchinson and Knopoff (1978, Appendix). For computing the sonance factor (Hofmann-Engl, 2004, 2008), the *Harmony Analyzer 3.2* applet software has been used, available at http://www.chameleongroup.org.uk/software/piano.html. For these models, always *C*4 (middle *C*) is taken as lowest tone.

chord class		emp. rank	roughness	instability	similarity	rel. periodicity	dual proc.
major	{0, 4, 7}	1 (1.667)	3 (0.1390)	1 (0.624)	1-2 (46.67%)	2 (4.0)	2
	{0, 3, 8}	5 (2.889)	9 (0.1873)	5 (0.814)	8-9 (37.78%)	3 (5.0)	1
	{0, 5, 9}	3 (2.741)	1 (0.1190)	4 (0.780)	5-6 (45.56%)	1 (3.0)	3
minor	{0, 3, 7}	2 (2.407)	4 (0.1479)	2 (0.744)	1-2 (46.67%)	4 (10.0)	4
	{0, 4, 9}	10 (3.593)	2 (0.1254)	3 (0.756)	5-6 (45.56%)	7 (12.0)	5
	{0, 5, 8}	8 (3.481)	7 (0.1712)	6 (0.838)	8-9 (37.78%)	10 (15.0)	6
susp.	{0, 5, 7}	7 (3.148)	11 (0.2280)	8 (1.175)	3-4 (46.30%)	5 (10.7)	7
	{0, 2, 7}	6 (3.111)	13 (0.2490)	11 (1.219)	3-4 (46.30%)	9 (14.3)	9
	{0, 5, 10}	4 (2.852)	6 (0.1549)	9 (1.190)	7 (42.96%)	6 (11.0)	8
dim.	{0, 3, 6}	12 (3.889)	12 (0.2303)	12 (1.431)	13 (32.70%)	12 (17.0)	12
	{0, 3, 9}	9 (3.519)	10 (0.2024)	7 (1.114)	10-11 (37.14%)	11 (15.3)	10
	{0, 6, 9}	11 (3.667)	8 (0.1834)	10 (1.196)	10-11 (37.14%)	8 (13.3)	11
augm.	{0, 4, 8}	13 (5.259)	5 (0.1490)	13 (1.998)	12 (36.67%)	13 (20.3)	13
correlation *r*			.352	.698	.802	.846	.791

Table 3: Consonance rankings of common triads. The empirical rank is adopted from Johnson-Laird et al. (2012, Experiment 1), where the tones are reduced to one octave in the theoretical analysis here. The roughness values are taken from Hutchinson and Knopoff (1979, Table 1), where again *C*4 (middle *C*) is taken as the lowest tone. For relative periodicity and percentage similarity (Gill and Purves, 2009), the frequency ratios from just tuning are used. The dual-process theory (Johnson-Laird et al., 2012, Figure 2) as a cognitive theory only provides ranks, no numerical raw values.

chord #	semitones	emp. rank		roughness		similarity		rel. periodicity		log. periodicity		dual proc.
1a (major)	{0, 16, 19}	1	(1.667)	2	(0.0727)	1	(64.44%)	1	(2.0)	1	(1.000)	1
2a	{0, 19, 22}	3	(2.481)	1	(0.0400)	5	(52.59%)	8-9	(12.3)	6-7	(3.133)	2
3a	{0, 16, 22}	4-5	(2.630)	3	(0.0760)	9	(38.62%)	15-16	(19.0)	8	(3.271)	3
4a	{0, 15, 22}	6	(2.926)	4	(0.0894)	8	(39.26%)	8-9	(12.3)	6-7	(3.133)	4
5a (minor)	{0, 15, 19}	2	(2.407)	5	(0.0972)	3-4	(55.56%)	4	(5.0)	4	(2.322)	5
6a (susp.)	{0, 7, 14}	7	(3.148)	6	(0.0983)	3-4	(55.56%)	6-7	(11.7)	5	(2.918)	6
7a	{0, 19, 23}	8	(3.370)	7	(0.1060)	2	(56.67%)	2-3	(4.0)	2-3	(2.000)	7
8a	{0, 16, 23}	4-5	(2.630)	8	(0.1097)	6	(52.22%)	2-3	(4.0)	2-3	(2.000)	8
9a (dim.)	{0, 15, 18}	11	(3.889)	16	(0.2214)	17	(28.57%)	13	(17.0)	14	(3.786)	9
10a	{0, 11, 14}	13	(3.963)	9	(0.1390)	18	(28.33%)	10-11	(14.3)	12	(3.585)	10
11a	{0, 18, 22}	15	(5.148)	12	(0.1746)	14	(30.05%)	15-16	(19.0)	11	(3.540)	11
12a	{0, 17, 23}	12	(3.926)	13	(0.1867)	12	(34.37%)	6-7	(11.7)	10	(3.497)	12
13a	{0, 14, 17}	9	(3.481)	14	(0.1902)	11	(36.11%)	5	(10.0)	9	(3.308)	13
14a	{0, 15, 26}	10	(3.630)	17	(0.2485)	13	(33.52%)	12	(15.7)	15	(3.800)	14
15a	{0, 11, 18}	14	(4.815)	18	(0.2639)	10	(36.90%)	17-18	(25.7)	17	(4.571)	15
16 (augm.)	{0, 16, 20}	16	(5.259)	10	(0.1607)	7	(41.67%)	10-11	(14.3)	13	(3.655)	16
17a	{0, 15, 23}	17-18	(5.593)	11	(0.1727)	16	(28.89%)	17-18	(25.7)	18	(4.655)	17
18a	{0, 20, 23}	19	(5.630)	15	(0.2164)	15	(29.44%)	14	(17.7)	16	(3.989)	18
19a	{0, 14, 25}	17-18	(5.593)	19	(0.3042)	19	(19.93%)	19	(101.0)	19	(5.964)	19
correlation r				.761	(.746)	.760	(.765)	.713	(.548)	.867	(.810)	.916
significance p				.0001	(.0001)	.0001	(.0001)	.0003	(.0075)	.0000	(.0000)	.0000

Table 4: Consonance correlation for complete list of triads in root position (Johnson-Laird et al., 2012, Experiment 2, $n = 19$). The ordinal rating and the numerical values of the considered measures are given in parentheses. The correlation and significance values written in parentheses refer to the ordinal rating, while the ones outside parentheses just compare the respective rankings. The roughness values are taken from Johnson-Laird et al. (2012, Figure 2), who employ the implementation available at http://www.uni-graz.at/richard.parncutt/computerprograms.html, based on the research reported in Bigand et al. (1996). In all other cases, just tuning is used as underlying tuning for the respective frequency ratios.

mode	semitones	emp. rank		sonance factor		similarity		log. periodicity (just tuning)		log. periodicity (rational tuning)	
Ionian	{0, 2, 4, 5, 7, 9, 11}	1	(0.83)	4	(0.147)	3	(39.61%)	1	(5.701)	1	(6.453)
Mixolydian	{0, 2, 4, 5, 7, 9, 10}	2	(0.64)	1.5	(0.162)	6	(38.59%)	4	(5.998)	3	(6.607)
Lydian	{0, 2, 4, 6, 7, 9, 11}	3	(0.58)	1.5	(0.162)	5	(38.95%)	2	(5.830)	2	(6.584)
Dorian	{0, 2, 3, 5, 7, 9, 10}	4	(0.40)	3	(0.152)	2	(39.99%)	3	(5.863)	4	(6.615)
Aeolian	{0, 2, 3, 5, 7, 8, 10}	5	(0.34)	6	(0.138)	4	(39.34%)	7	(6.158)	5	(6.767)
Phrygian	{0, 1, 3, 5, 7, 8, 10}	6	(0.21)	7	(0.126)	1	(40.39%)	5	(6.023)	6	(6.778)
Locrian	{0, 1, 3, 5, 6, 8, 10}	7		5	(0.142)	7	(37.68%)	6	(6.033)	7	(6.790)
correlation r				.667		.036		.786		.964	
significance p				.0510		.4697		.0181		.0002	

Table 5: Rankings of common heptatonic scales (church modes), i.e. with 7 out of 12 tones. As empirical rating, the overall preference for the classical church modes is adopted (Temperley and Tan, 2013, Figure 10). For the sonance factor (Hofmann-Engl, 2004, 2008), again $C4$ (middle C) is taken as lowest tone. For percentage similarity, the values are taken directly from (Gill and Purves, 2009, Table 3).