THE INFLUENCE OF PITCH ON MELODIC SEGMENTATION

Tillman Weyde
University of Osnabrück
Research Department of Music and Media Technology
tweyde@uos.de

ABSTRACT

Melodic segmentation is an important topic for music information retrieval, because it divides melodies into musically relevant units. Most influential theories on melodic segmentation of the last decades have stressed the role of pitch for melodic segmentation. The general assumption was, that relatively large changes or distances in any musical parameter like pitch, time, dynamics, or melodic movement mark segment boundaries. This has generally been accepted despite the lack of empirical studies. Here an empirical study is presented, that investigates the influence of inter-onset-intervals (IOI), intensity accents, pitch intervals, and pitch interval direction changes. The results show a significant influence only for IOIs and intensity, but neither for pitch intervals nor for changes in interval direction. The validity of the results and possible explanations are discussed and directions of further investigations are outlined.

1 Introduction

The segmentation of melodies is an essential part of melody perception and cognition. Melodic segments or motifs form the basic elements of melodic structure, like words are the basic meaningful elements of speech. The importance of segmentation has long been discussed by theorists (e.g. Riemann, 1884), and it has been the subject of theories (Lerdahl and Jackendoff, 1983, Cooper and Meyer, 1960) and computer models (Tenney and Polansky, 1980, Cambouropoulos, 2001, Temperley, 2001). As melody comprises pitch, rhythm, dynamics, and implicitly harmony, most approaches tried to identify influential factors in all of these domains. The general assumption was that the Gestalt rules of proximity and similarity are the basis for the introduction of boundaries between musical phrases. There are of course other factors, like the recognition of known patterns, closure or *Prägnanz*, but their influence has less often been investigated (see Bod, 2001, Weyde, 2002). The idea of treating the different dimensions elegantly by applying the same principle has lead to some questionable hypotheses.

There have been very few empirical studies, mainly to evaluate existing segmentation algorithms against the examples of musical segmentations (see Spevak et al., 2002). The basic idea of this work is to systematically test the effect of different musical parameters on the perceived segmentation. For these tests synthetic melodies have been generated, which vary musical parameters independently and systematically.

2 Experimental Design

The experiment has a forced choice design, asking subjects for the length of segments when listening to a melody. Subjects were presented short melodic sequences, which were designed to be completely isochronous and uniform, apart from two conflicting segmentation cues, of which one indicated a segmentation into groups of two notes and the other into groups of three.

Four factors were tested: inter-onset-intervals, loudness accents, pitch intervals, and changes in pitch direction. Some authors suggest to use the intensity differences instead of accents Lerdahl and Jackendoff (1983), but this approach was not pursued because already Woodrow (1909) had found in his experiments, that relatively loud notes tend to mark the beginning of a new group, while soft ones do not.

Each of the factors was varied in five steps, resulting in 25 listening samples per pair of factors, and all six possible pairs of factors were tested resulting in 150 listening examples. This approach of testing pairs of factors was chosen to reduce the number of samples in the experiment, which would have been 650 for all possible combinations of the four factors in five steps.

The experiments were conducted in two passes, each covering three of the six pairs. The stimuli were presented via MIDI with a piano sound on a personal computer with a program, that allowed the subjects to repeat the playback and to make a choice of preferred segmentation into groups or 2 or 3. The melodies comprised twelve notes and were presented at a comfortable volume level. The subjects could listen as often and long as they wanted, but they had to make a choice. After the user made a choice, a break of 3 seconds was introduced to avoid integrated perception of the melodies. Within each factor-pair the stimuli were presented in random order. In the first pass eight subjects participated, and nine in the second set. All subjects were music students between 21 and 26 years of age.

3 Results

In the following, the stimuli and results and a regression analysis of the six factor-pairs are presented.

Figure 1. Piano-Roll of an Ioi/Velocity stimulus. Darker rectangles depict louder notes. The background lattice has one horizontal line per semitone and one vertical line per second.

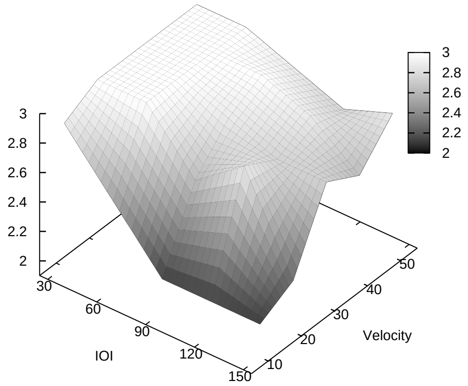

Figure 2. Average choice of segment length in the IOI-Velocity experiment.

3.1 Ioi/Velocity

In this set of stimuli, intensity accents were applied every three notes. Intensity was realized by using MIDI-Velocity, the unaccented value was 70 and the accent range was from 10 to 50. By subjective examination the sound generator responded approximately linear to the velocity values in terms of perceived loudness. Longer IOIs were inserted before every third note in the range from 30 to 150 milliseconds. A visualization is shown in figure 1.

The average response of the subjects is shown in figure 2. The responses were as could be expected for a categorization task, showing sensitivity to both factors and saturation effects for high velocity and low IOIs.

A logistic regression was performed on the experimental data with the following results shown in table 1. Both IOI and velocity have significant influence on the choice and the logistic model accounts for 33% of the variance.

3.2 Ioi/Pitch

Here pitch changes were inserted every three notes, alternating upward and downwards. The intervals were in the range form 1 to 5 starting with middle c. The IOIs were varied like in the last experiment. A visualization is shown in figure 3.

The average response of the subjects is shown in figure 4. It shows clearly that the effect of IOIs is much stronger than that of pitch. This is also confirmed by the logistic regression results, showing that IOIs have significant influence (see table 2).

Model Log Likelihood	-74.557
Intercept Log Likelihood	-111.355
R^2	.330

	Coef	Std. Err	Coef/SE	χ^2	P-Value
const.	1.540	.642	2.399	5.755	.0164
ioi	$-.031$	.006	-5.176	26.790	$< .0001$
velo	.102	.018	5.509	30.349	$< .0001$

Table 1. Logistic regression of IOI/Velocity

Figure 3. Piano-Roll of an Ioi/Pitch stimulus.

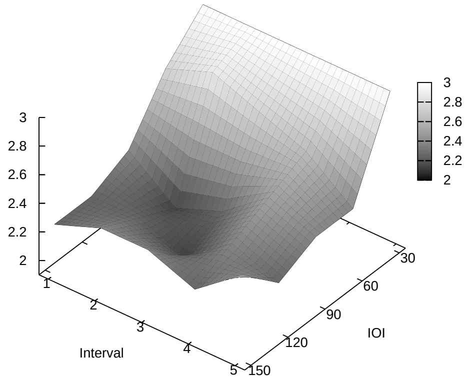

Figure 4. Average choice of segment length in the IOI-Pitch experiment.

3.3 Ioi/Direction

In these stimuli there was a pitch change after every note which changed direction every three notes, alternating upward and downwards. The intervals were in the range form 1 to 5 starting with middle c. The IOIs were varied like in the last set. A visualization is shown in figure 5.

The average response (figure 6) shows predominant effect of IOIs. For an interval of 2 semitones, the direction change has no effect at all. The logistic regression (table 3) shows that again only IOIs have significant effect.

3.4 Pitch/Direction

These stimuli combine the regular intervals changing direction every three notes with additional intervals every two notes. Both the regular and the additional intervals were varied from 1 to 5 semitones (see figure 7).

The average response (figure 8) shows no clear tendency. In the logistic regression (table 4) the influence of the additional intervals is stronger than that of the direction change, but neither has significant effect.

3.5 Pitch/Velocity

These stimuli combine the regular intervals changing velocity every three notes with additional intervals every two notes. Both the regular and the additional intervals were varied from 1 to 5 semitones (see figure 9).

Model Log Likelihood	-112.705
Intercept Log Likelihood	-138.469
R^2	.186

	Coef.	Std. Err	Coef/SE	χ^2	P-Value
const.	2.584	.549	4.711	22.190	$< .0001$
ioi	-.027	.004	-6.343	40.231	$< .0001$
pitch	-.013	.115	-.115	.013	.9086

Table 2. Logistic regression IOI/Pitch

Figure 5. Piano-Roll of an Ioi/Direction stimulus.

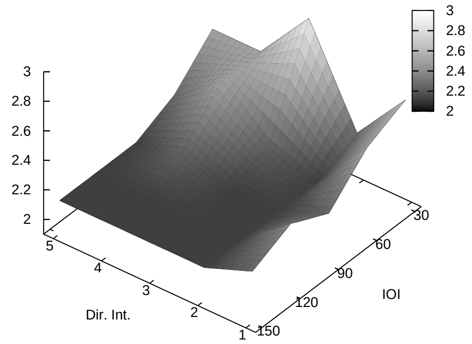

Figure 6. Average choice of segment length in the IOI-Direction experiment.

The average response (figure 10) shows no clear tendency. In the logistic regression (table 5) the influence of the additional intervals is stronger than that of the velocity change, but neither has significant effect.

3.6 Direction/Velocity

These stimuli combine the regular intervals changing velocity every three notes with intensity accents every two notes as in the previous examples. A sample is shown in figure 11.

The average response (figure 12) shows no clear tendency. In the logistic regression (table 6) the influence of the additional intervals is stronger than that of the velocity change, but neither has significant effect.

4 Discussion

The results of the pairwise tests show significant effects for velocity and IOIs, but not for pitch or change of pitch interval direction. This result is consistent over all pairs of factors. This contrasts to most currently accepted theories on melodic perception. There are several possible explanations for this effect, with different consequences.

4.1 Experimental Design Considerations

In the results the level of variance accounted for by the factors in the model is not very high. So there may be other factors, that influence the decision. One possibility is that the pauses between the examples did not suffice to prevent the establishing of a metrical structure in the listener that could bias the perception of the next sample. The perception of melodic segmentation can to some

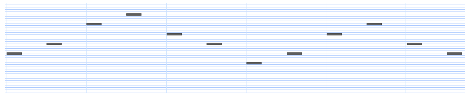

Figure 7. Piano-Roll of a Pitch/Direction stimulus.

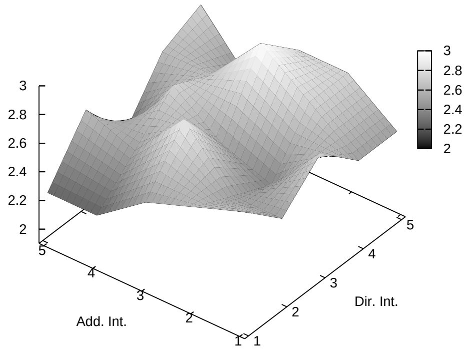

Figure 8. Average choice of segment length in the Pitch/Direction experiment.

degree be controlled by the listener, therefore personal or momentary preferences may account for a part of the variance. There are obviously other factors like harmony and tempo which can have an influence on the segmentation. Most important to mention here is the recognition of high-level structures (e.g. repetitions, harmonic progression) and the recognition of patterns known form other contexts.

Another argument against the validity of the experimental results could be, that the stimuli were synthetic, and that perception may be different in the context of real music. Evaluating segmentation in the context of real music is difficult, because of the interaction of multiple factors and measuring the similarity of segmentations (see Spevak et al., 2002). On the other hand, if the rules like the *Grouping Preference Rules* of Lerdahl and Jackendoff (1983) were general laws of perception, they should have effect in synthetic settings, too.

4.2 The Role of Pitch in Melodic Segmentation

After all, the results are consistent throughout the tests, showing significant influence of pitch and velocity. It seems therefore justified to conclude that rhythm and dynamics have a considerably stronger influence on segmentation than pitch interval size and direction changes. At least considering linear and log linear effects in the ranges tested here.

It seems plausible that pitch does have a significant effect on melodic segmentation, since it plays such an important role in melody and music in general. It is possible that only large intervals have an effect and that the range

Model Log Likelihood	-120.305
Intercept Log Likelihood	-132.459
R^2	.092

	Coef.	Std. Err	Coef/SE	χ^2	P-Value
constant	.842	.483	1.742	3.036	.0814
Dir	-.100	.112	-.889	.791	.3737
IOI	-.018	.004	-4.549	20.697	$< .0001$

Table 3. Logistic regression IOI/Direction

Model Log Likelihood	-133.649
Intercept Log Likelihood	-135.725
R^2	.015

	Coef.	Std. Err	Coef/SE	χ^2	P-Value
constant	.226	.455	.497	.247	.6194
int	.167	.103	1.623	2.635	.1046
dir	-.126	.103	-1.223	1.496	.2213

Table 4. Logistic regression Pitch/Direction

Figure 9. Piano-Roll of a Pitch/Velocity stimulus.

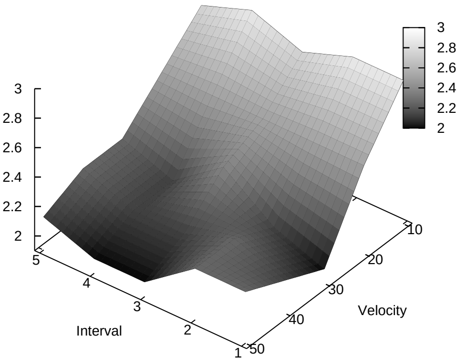

Figure 10. Average choice of segment length in the Pitch/Velocity experiment.

Figure 11. Piano-Roll of a Direction/Velocity stimulus.

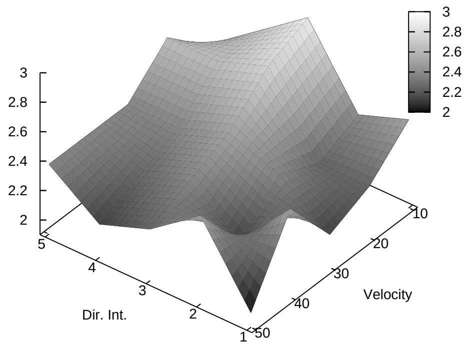

Figure 12. Average choice of segment length in the Direction/Velocity experiment.

of the intervals used were too small. Another possibility is that there are more complex effects which are not captured by the regression analyses.

To bring more certainty into this issue, it would be useful, to conduct further experiments. The current experiment does have some shortcoming, as it was not designed to prove the general significance of pitch to melodic segmentation. To make sure that there is no effect from one presentation to other, one could play some noise inbetween. It is also necessary to have examples which have only one segmentation factor, and to vary the length of the segmentation. These extension of the experiment will require a larger number of tests, with larger numbers of subjects.

5 Conclusion

The results of the experiment described here indicate that the size of pitch intervals have little influence on melodic segmentation compared to timing and dynamics. It this were confirmed by more data, it would have important consequences for music information retrieval. In that case, segmentation algorithms could largely ignore pitch interval size. Instead other models on how pitch influences melodic segmentation would have to be evaluated. For establishing certainty on this question, further experiments are needed.

References

Bod, R. (2001). Memory-based models of melodic analysis: Challenging the gestalt principles. *Journal of New Music Research*, 30(3).

Cambouropoulos, E. (2001). The local boundary detection model (lbdm) and its application in the study of expressive timing. In *Proceedings of the International Computer Music Conference 2001*, pages 290–293, Havanna, Cuba.

Cooper, G. W. and Meyer, L. B. (1960). *The Rhythmic Structure of Music*. University of Chicago.

Lerdahl, F. and Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. The MIT Press, Cambridge, Mass.

Riemann, H. (1884). *Musikalische Dynamik und Agogik*. Hamburg, Leipzig, St. Petersburg.

Spevak, C., Thom, B., and Höthker, K. (2002). Evaluating melodic segmentation. In Anagnostopoulou, C., Ferrand, M., and Smaill, A., editors, *Music and Artificial Intelligence. Proceedings of the Second International Conference on Music and AI*, volume 2445 of *Lecture Notes on Artificial Inetelligence*, pages 168–182. Springer-Verlag.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. The MIT Press, Cambridge, Mass.

Tenney, J. and Polansky, L. (1980). Temporal gestalt perception in music. *Journal of Music Theory*, 24(2):205–241.

Weyde, T. (2002). Integrating segmentation and similarity in melodic analysis. In Stevens, K., Burnham, D., McPherson, G., Schubert, E., and Renwick, J., editors, *Proceedings of the International Conference on Music Perception and Cognition 2002*, pages 240–243, Sydney, Australia. University of New South Wales.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology*, 14:1–66.

Model Log Likelihood	-115.836
Intercept Log Likelihood	-148.659
R^2	.221

	Coef.	Std. Err	Coef/SE	χ^2	P-Value
constant	2,247	,522	4,307	18,551	$<,0001$
Pitch	-,065	,114	-,570	,325	,5689
Velo	-,092	,013	-6,910	47,749	$<,0001$

Table 5. Logistic regression Pitch/Velocity

Model Log Likelihood	-144.789
Intercept Log Likelihood	-151.824
R^2	.046

	Coef.	Std. Err	Coef/SE	χ^2	P-Value
constant	.215	.435	.493	.244	.6217
Dir	.137	.100	1.378	1.899	.1681
Velo	-.035	.010	-3.413	11.650	.0006

Table 6. Logistic regression Direction/Velocity