PROSODIC RISE AND RISE-FALL CONTOURS AND MUSICAL RISING TWO-TONE PATTERNS

Ernst Dombrowski, Thurid Holzrichter, Niels Münz, Alexander Nowak, Monika Poschmann

Department of Psychology, Christian-Albrechts-University Kiel
ed@ipds.uni-kiel.de

ABSTRACT

A series of musical rising two-tone patterns comprising 13 intervals, from unison to octave, is compared with two German nuclear accents, *rise* and *rise-fall*, both on disyllabic utterances with upbeat-downbeat structure. The two-tone patterns are viewed as potential musical representatives of the prosodic contours. However, because of rhythmic constraints they can only cover part of the prosodic contour information. Two ABX tests were carried out to examine whether the loss of contour information in the two-tone patterns is compensated for with other musical features. Results show that large and dissonant intervals are more likely assigned to the *rise* contour. Moreover, the *rise* is judged closer to the rising fifth (a I-V progression) and the *rise-fall* closer to the rising fourth (a V-I progression). This means that the relation between speech melody and musical melody is not restricted to melodic contour but includes aspects that are usually treated as exclusively musical. Thus, the loss of contour information in the musical patterns should not be interpreted in terms of contour truncation but in terms of a feature replacement applying to the phrase-final features of melodic patterns.

Keywords: prosody and music, melody.

1. INTRODUCTION

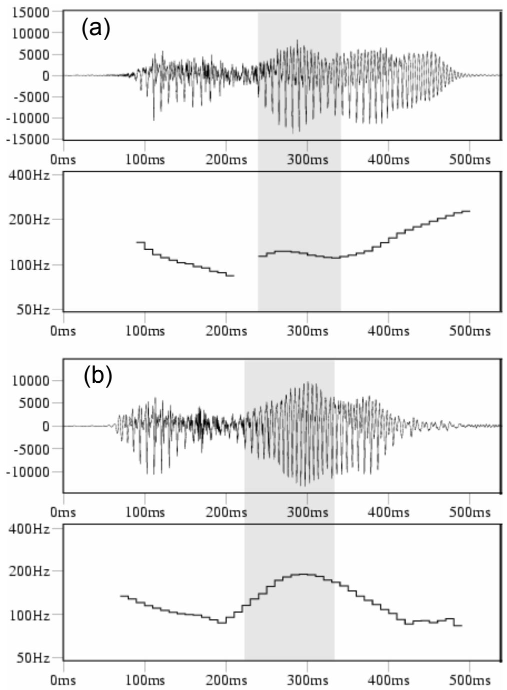
According to Lerdahl [9], the analogy between speech and music is based on prosodic factors, mainly rhythm and melodic contour. Tonal structure (in the sense of pitches, intervals, tonality, harmony etc.) is excluded from the parallels. However, for the relations between the melodic contours of speech and music this point of view would mean that they can only be outward – because an essential aspect in the formation of musical melodic patterns is omitted. Therefore, an interesting issue is whether similarities between speech melody and musical melody are actually restricted to contour information or whether they include other features, in particular features that are usually declared to be genuinely musical.

The effect of non-contour information on the perceived relation between the melodic patterns of speech and music can be revealed if a part of the prosodic pitch movement is lost in the transition from the plain speech mode to the stylized musical mode of melody. Losses of this kind are due to the transformation of continuous contours into stepped contours. They are favoured by rhythmic restrictions, namely if there are too few rhythmic events to represent all prosodic pitch movements in musical tonal steps.

Figure 1: The first triplet of *Wohin* calls from the bass aria No. 24 in J. S. Bach's St John Passion BWV245, bars 69 to 51 [1].

This can be shown with an example, the *Wohin* calls in the Bass Aria No. 24 *Eilt, ihr angefochtnen Seelen* ("Come, ye souls whom care oppresses") from J. S. Bach's St John Passion, BWV 245 [1, 5]: In this aria, sets of short calls of the choir with the German text *Wohin* ("Where") are inserted in the bass solo. The calls are set to music with rising two-tone patterns in various musical intervals (Fig. 1). For a 'retranslation' of the rising calls into speech melody, based on German intonation, two patterns seem to suit best: *rise* and *rise-fall* (cf. Fig. 2a and b, [5]). These contours are partly similar, partly different. Both begin with a rising movement towards the accented-vowel. However, in one case this initial rise is low leaving room for the rising movement to continue. In the other case the initial rise is high, facilitating the ensuing fall.

Figure 2: The German utterance *Wohin* [voɦˈɪn] produced with a *rise* (a) and a *rise-fall* (b). Oscillogram, F0 contour. The accented vowel [ɪ] is highlighted in grey.

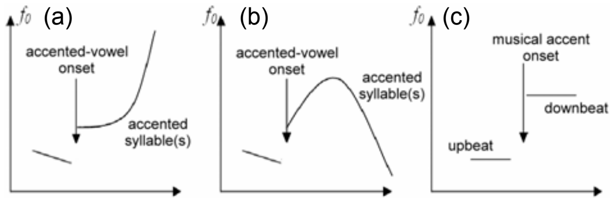


Thus, the difference between the two contours can be understood in terms of terminal vs. non-terminal modifications after a similar beginning, i.e. an initial rising step.

The contour characteristics emphasized here are reflected in transcriptions using current models of German intonation: In the AM perspective of German ToBI [7], the two patterns are an LH* or L*H pitch accent with high boundary tones and an LH* pitch accent with low boundary tones. According to the Kiel Intonation Model (KIM) [8], they are a high rising (i.e. interrogative) early valley and a falling (i.e. terminal) medial peak. The musical counterpart of these contours only contains their rising part (cf. the schematic representation in Fig. 3).

The resulting ambiguity of the mutual assignments of *two* prosodic patterns and *one* musical pattern originates from the stepped musical contour shape combined with a restriction of the number of rhythmic events (i.e. syllables and tones) to only two in one accent group. Only one of the original melodic components is preserved in the musical contour, i.e. in the up and down of pitches. Nonetheless, the musical version sounds complete.

Figure 3: Schematic F0 shapes of a *rise* (a), a *rise-fall* (b), and a rising musical two-tone pattern (*stylized rise*) (c). Vertical arrows indicate the accented-vowel onset and the musical accent onset. The rhythmic structure is upbeat-downbeat.



Starting with this observation, two perception experiments have been carried out to examine the question whether there are musical means that can compensate for the loss of contour information if a prosodic pattern is transformed into a corresponding musical one. These musical means would establish a clear relation between *one* prosodic and *one* musical pattern. Analogous to the *Wohin* calls, the experiments are based on upbeat-downbeat sequences consisting of two events: syllables or tones. The rising two-tone pattern is realized with musical successive intervals of different size, from the unison to the octave, using instrumental instead of vocal sounds. In this way, three aspects can be analysed that could modify the perception of rising patterns: (1) The *interval size*, i.e. the *pitch range* of an interval, which is a further contour feature beyond the 'up and down'. In a rising movement, the pitch range may express different degrees of tension or openness. (2) The *pitch relationship* of the successive intervals, their *musical consonance*. Dissonance in a rising movement can be seen as an indicator of progression as opposed to non-progression and again of tension as opposed to relaxation (cf. Lerdahl and Jackendoff [10]). (3) *Inferred musical syntax information* or *function*, reaching or leaving the tonic. Here, the rising *fourth* and the rising *fifth* are compared (most likely) suggesting a V-I vs. I-V progression, i.e. a movement from the dominant towards the tonic and vice versa. German *rise* and *rise-fall* contours are used as prosodic comparative patterns, both with a rising step towards the accented vowel.

2. METHOD

Experiment 1 was an *ABX interval matching task*: Listeners assessed on a five-point scale whether a rising two-tone pattern came closer to a *rise* or to a *rise-fall*. The two-tone patterns were presented with 13 musical intervals, from unison to octave. For the prosodic patterns the German disyllabic utterance *Für S'ie/s'ie* ("for you/her") was used as a carrier.

Experiment 2 was an *ABX contour matching task*: Here, listeners assessed on a five-point scale whether a spoken contour fitted a rising fourth or fifth. The prosodic patterns, *rise* and *rise-fall*, were presented on four disyllabic utterances: *Für S'ie/s'ie*, *Mal s'ehn* ("do you want to have a look"), *Wies'o* ("why"), *War'um* ("why").

Stimuli: The four test utterances with the contours *rise* and *rise-fall* were produced by a trained Northern German speaker. Both contours were spoken with the pragmatic background 'question'. For the upbeat syllables a pitch of an A2 (ca. 110 Hz) was taken as a reference and for both contours equal pitch maxima were intended. The averaged actual F0 maximum was 236 Hz for the *rises* and 181 Hz for the *rise-falls*. Based on the natural utterances contours with adapted F0 maxima were resynthesized using PRAAT. Since the post-processed *rise* contours, however, lost the pragmatic character of neutral questions the natural utterances were included in the experiments. As an additional task following experiment 1 and 2, listeners (see below) compared the natural and the post-processed contours with respect to their perceived height. In the natural condition, the *rises* got more *higher*-answers than the *rise-falls*. The post-processed contours differed as well, but in the opposite direction, which additionally justified the inclusion of the natural stimuli in the experiments.

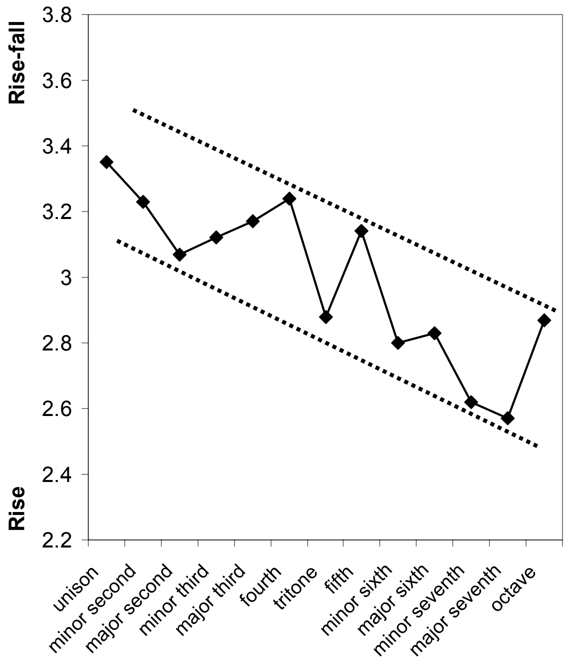
The musical two-tone patterns were synthesized as piano sounds (with *Cubase*). The durations of the tones were approximated to duration proportions calculated from the voiced parts of the test utterances. However, the total duration of the musical patterns was twice as long – in accord with the duration relation usually observed between sung and spoken speech. Three randomized series were produced for each experiment, one using A2 (110 Hz) as a starting point, the other two were transpositions using F#2 and C3. The first tone was the same in all intervals of one series, the second varied. Sound examples are available in [5].

Subjects: 31 musically trained listeners participated in the study (24 females, 7 males, aged from 19 to 26 years, most of them were students of psychology). Musical training was ascertained with a self-assessment questionnaire.

3. RESULTS

A repeated measures ANOVA on experiment 1 yields significant differences between the 13 musical intervals (p=0.002). As shown in Fig. 4, listeners tend to assign large rising intervals to the *rise* and small rising intervals to the *rise-fall*. However, there is no linear increase of the *rise* preference with interval size. Instead, on the increase, effects of musical consonance vs. dissonance are superimposed: The rising fourth and the rising fifth are judged to be closer to the *rise-fall* pattern than the neighbouring tritone. The same holds for the octave compared with the minor and major sevenths. In the graph of the means (Fig. 4), the effect of musical consonance is illustrated by dotted lines connecting the means of the consonant intervals and the means of the dissonant or less stable intervals respectively.

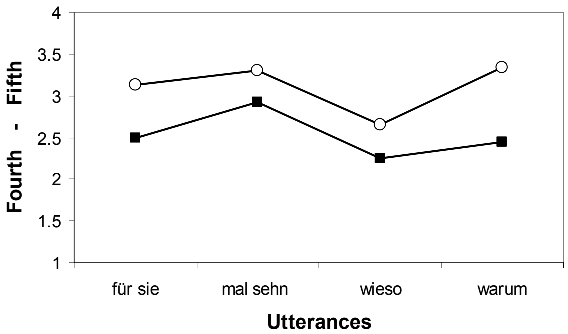
Figure 4: Averaged assignments of 13 musical successive intervals (unison to octave) to prosodic *rise* vs. *rise-fall* contours on the German utterance *Für Sie/sie*. Judgements on a 5-point scale. The dotted lines link the means observed for consonant intervals and for dissonant or less stable intervals respectively.



In addition, there is a significant effect of presentation order (p=0.001) and an interaction of interval size and presentation order (p=0.037): i.e. listeners are also inclined to assign intervals to the prosodic pattern they heard last, which is in particular true for large intervals (except for the octave). Therefore, if the heard order is *rise – rise-fall* the consonance effect is more prominent, in the reverse order the interval-size effect prevails. There are no significant differences between the transposition conditions.

A repeated measures ANOVA on experiment 2 shows for all utterances that *rises* and *rise-falls* differ in their affinity to the rising fifth vs. fourth intervals (p=0.006, p=0.001, p=0.023 for the three transposition conditions). *Rises* are preferably related to the fifth, *rise-falls* to the fourth (with the fifth suggesting a I-V progression, the fourth suggesting a V-I progression, i.e. a half cadence vs. a full cadence); cf. Fig. 5.

Figure 5: Averaged assignments of four disyllabic utterances, *rise-fall* = **O** vs. *rise* = ■, to musical rising fifths vs. fourth intervals. A 5-point scale was used.



4. CONCLUSIONS

According to the results, the mutual assignment of the musical and prosodic patterns under study is determined by a set of three factors that compensate for the presumed loss of contour information in musical rising two-tone configurations (e.g. in the above Bach example). One factor is another contour feature: *interval size*. Two are based on non-contour information: *musical consonance* and *implied tonal functions*. Thus, the data indicate a systematic relation between prosodic and musical features that goes beyond rhythm and contour similarity and that includes musical pitch.

The prosodic and musical patterns analysed in the experiments may be seen in a common frame (also cf. [3, 4]): (1) The rising two-tone patterns can be interpreted in terms of current cognitive music theory. They are basic musical units expressing *tension* vs. *relaxation* or at least different degrees of musical *tension*, concepts used by Lerdahl and Jackendoff [10] to explain structural coherence in music. Moreover, the rising patterns can be related to different types of musical closure (*half cadence* vs. *full cadence*). (2) A concept of tension vs. relaxation may also be applied to the two prosodic patterns and can thus be a mediator between the domains of music and prosody. Furthermore, *rise* and *rise-fall* patterns can be viewed as patterns with different final contour modifications, i.e. again with different types of closure.

The two above experiments support structural parallels of this kind showing that features of musical pitch may stand in for prosodic contour features (and vice versa) to signal tension or closure. This perspective of the relationship between melodic patterns in speech and music implies a particular formal description: If there is a loss of contour in the transformation of a structural unit of speech melody into a corresponding musical unit this should not be viewed as merely cutting off a contour section (i.e. *truncation*; cf. [6, 2]) but as a replacement of contour information by other suitable information (i.e. *feature replacement*).

5. ACKNOWLEDGEMENTS

We thank the IPDS, University of Kiel, for providing the facilities to construct the sound materials for the experiments.

6. REFERENCES

[1] Bach, J. S. (1981). St John Passion: BWV 245, Vocal Score (Supplementary edition based on J. S. Bach, *Neue Ausgabe sämtlicher Werke*, II, 4). Kassel: Bärenreiter.

[2] Chen, M. Y. (1983). Towards a grammar of Singing: tune-text association in Gregorian Chant. *Music Perception*, 1, 84-122.

[3] Dombrowski, E. (1995). Über strukturelle Gemeinsamkeiten zwischen sprachlichen und musikalischen Melodien. In: Behne, K.-E., de la Motte-Haber, H., Kleinen, G. (eds), *Jahrbuch der deutschen Gesellschaft für Musikpsychologie*. Vol. 12. Wilhelmshaven: Noetzel, 110-123.

[4] Dombrowski, E. (2003). Steps to a common description of melody in speech and music. *Proc. of the 5th Triennial ESCOM Conference*. Hanover: University of Music and Drama, 479-483.

[5] Dombrowski, E., Holzrichter, T., Münz, N., Nowak, A., Poschmann, M. (2007). *Prosody and musical rising two-tone patterns: sound examples*. URL http:// www.ipds.uni-kiel.de/publikationen/publikationen.de.html

[6] Grabe, E. (1998). Pitch-accent realization in English and German. *J. Phonet.* 26, 129-144.

[7] Grice, M., und Baumann, S. (2000). Deutsche Intonation und GToBI. *Linguistische Berichte*, 191, 267-298.

[8] Kohler, K. (1991). A model of German intonation. *Arbeitsberichte des Instituts für Phonetik der Universität Kiel (AIPUK)*, 25. Kiel: IPDS, 295-368.

[9] Lerdahl, F. (2001). The sounds of poetry viewed as music. In: R. J. Zatorre and I. Peretz (Eds.), *The biological foundations of music*. New York: The New York Academy of Sciences, 337-354.

[10] Lerdahl, F., and Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.