A Formal Theory for the Discovery of Local Boundaries in a Melodic Surface

Emilios Cambouropoulos, Faculty of Music & Department of Artificial Intelligence,
University of Edinburgh, 12 Nicolson square, EH8 9DF Edinburgh, emilios@music.ed.ac.uk

Abstract: *In this paper a general theory will be introduced that facilitates the detection of local boundaries in melodic surfaces. A model will be proposed that discovers points of maximum local change that allow a listener to identify potential perceptual boundaries. It is suggested that the Local Boundary Detection Model (LBDM) presents a more effective method for low-level segmentation in relation to other existing models and it may be incorporated as a supplementary module to more general grouping structure theories.*

1. Introduction

The Gestalt principles of perceptual organisation are a set of rules-of-thumb that suggest preferential ways of grouping mainly visual events into larger scale schemata. [Tenney 1961] discusses the use of the principles of proximity and similarity as a means of providing cohesion and segregation in 20th century music and, later, [Tenney & Polansky 1981] develop a computational system that discovers grouping boundaries in a melodic surface. Musical psychologists [see Deutsch 1982a,b; McAdams 1984; Bregman 1990] have suggested and experimented as to how the Gestalt rules may be applied into auditory/musical perception and [Deutsch & Feroe 1981] further incorporate such rules in a formal model for representing tonal pitch sequences. The grouping component of the Generative Theory of Tonal Music (GTTM) [Lerdahl & Jackendoff 1983] is based on the Gestalt theory and an explicit set of rules is thereby described - especially for the low-level grouping boundaries (the formulation of these rules has been supported by the experimental work of [Deliège 1987]).

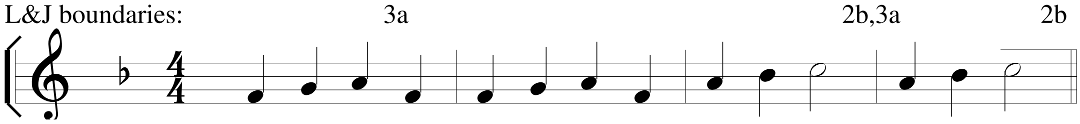
In this paper a systematic theory will be described that attempts to define local boundaries in a given melodic surface. The proposed segmentation model (*Local Boundary Detection Model - LBDM*) will be based on two rules: the *Identity-Change rule* (which is lower-level than the Gestalt principles of proximity and similarity) and the *Proximity rule* (which is an extended version of the Gestalt proximity principle). It will be shown that the formulation of the boundary discovery procedures defined by Lerdahl & Jackendoff and Tenney & Polansky are merely specific instances of the proposed theory. Most of the examples and counter-examples will be given in relation to the influential formulation of the local detail grouping preference rules (mainly GPR 2 & 3) and supporting examples presented by [Lerdahl & Jackendoff 1983].

More specifically, the following issues will be addressed:
1. What are some problems, limitations and inconsistencies in the way the Gestalt rules of proximity and similarity are applied to temporal musical sequences by existing theories? How might these problems be resolved?
2. Is there a more general and clear way to define the low-level rules of perceptual organisation? How may the existing rules relate to such lower-level principles?
3. Can one construct a general formal theory that suggests *all* possible grouping boundaries rather than a theory that relies on a restricted set of heuristic rules?
4. How can this general fundamental theory be further refined?

Our goal is to develop a formal theory that may suggest all the possible points for local grouping boundaries on a musical surface with various degrees of prominence attached to them. These are only seen as *potential* boundaries as one has to bear in mind that musically interesting groups can be defined only in conjunction with higher-level grouping analysis (parallelism, symmetry etc.). Low-level grouping boundaries may be coupled with higher-level formations so as to produce 'optimum' segmentations (see fig. 1).

figure 1: Beginning of *Frère Jacques*. Higher-level grouping principles override some of the local detail grouping boundaries (note that *LBDM* gives local values at the boundaries suggested by parallelism - without taking in account articulation - whereas Lerdahl and Jackendoff do so only for the 3rd and 4th boundary).

L&J boundaries: 3a 2b,3a 2b



LBDM boundaries: 0 1 **4** 2 2 1 **2** **2** 1 4 **10** 2 5 **8**

parallelism:

2. Problems in the application of the low-level Gestalt rules in existing theories

The low-level Gestalt principles of perceptual - mainly visual - organisation state that objects closer together (Proximity principle) or more similar to eachother (Similarity principle) tend to be perceived as groups. Some problems in the way these rules have been applied in the organisation of temporal musical sequences are briefly discussed below:

1. [Tenney and Polansky 1981] state that the similarity principle - as they define it - actually includes the proximity principle as a special case. "In both, it is the occurrence of a *local maximum in interval magnitudes* which determines clang-initiation" (p. 211). In the GTTM [Lerdahl & Jackendoff, 1983] the grouping rules are defined in such a way that it seems rather plausible that the proximity rules can be subsumed by the change (similarity) rules and the reverse. For example, GPR3a (register rule) states that a greater pitch interval inbetween smaller neighbouring intervals initiates a grouping boundary. This can been seen in two ways: a) that the pitches of the first and last intervals are more *similar* to each other than the pitches of the middle interval or b) that there is a greater *proximity* between the first two pitches (and the last two) rather than between middle pitches. [Handel 1989, p.198] states: "Proximity really refers to spatial proximity and when we say that two frequencies are near each other we are really employing a spatial analogy. We could just as easily have said that they were more similar to one another. However, it really does not matter since the effects of the two are the same. Things that are either more similar to one another than to their neighbours, or nearer to one another, will tend to be grouped together perceptually."

As defined in these theories the proximity and similarity principles are merely different descriptions of the same phenomenon, namely a local maximum in interval magnitudes. In the current paper it will be maintained that although this definition provides the most important factor for discovering local boundaries a more general approach should also account for local *minima* (actually for any change) in interval magnitudes. For example, in the following sequence of durations: ♩ ♩ ♩ ♪ ♩ ♩ ♩ a listener easily hears a possible point of segmentation (neither the Tenney & Polansky nor the Lerdahl & Jackendoff formalisms suggest any boundary). For this reason a different, more elementary rule will be introduced based on the principle of identity/change. This issue will be discussed further in the following sections and it will be shown that the above example can naturally be accommodated within *LBDM*.

2. The way that the Gestalt principles are usually defined, implies that the proximity principle applies for spatial (or temporal) intervals/distances between different objects whereas the similarity principle applies for internal attributes of objects. It seems though natural - at least for music - to relate the similarity principle to intervals as well (musical intervals as defined by [Lewin 1981]). Musical intervals may be considered 'objects' themselves perceived by the musical mind. This is rather obvious for the domain of musical time, as durations are simultaneously attributes of individual musical events and intervals between time points of different events. But pitch and loudness intervals (timbre intervals are not discussed in the current paper) can also be perceived as objects independent of actual pitch and loudness values. Pitch intervals tend to be perceived categorically [Handel, 1989, p.281-283] and play a leading role in the perception and memorisation of melodic schemata [Dowling, 1986, ch. 5]. Loudness has a built-in interval aspect as even in the physical domain a decibel - for example - is defined as "a unit of difference between two sounds in terms of the ratio between sound pressures" [Handel, 1989, p.551]. The assumption made in the current proposed model is that the similarity principle can be directly applied to musical intervals. The proximity principle is also extended so that it may be applied to intervals between any parametric attributes of musical events (not only temporal distances).

3. The underlying assumption of any attempt to apply the principles of visual perceptual organisation into the temporal domain is that time can be modelled as a one-dimensional space (see [Xenakis 1992] on outside-time and in-time aspects of music). Musical scores are possible because of this sort of mapping.

There is though one fundamental difference that is often overlooked or is implicitly presupposed that makes this mapping not as obvious as it initially seems. This has to do with the fact that temporal objects are not simply mapped on a one-dimensional space but onto a *directional* one-dimensional space i.e. an axis with a given positive and negative direction (in other words, we can tell left from right). The low-level Gestalt principles of proximity and similarity are usually applied on symmetrical spaces. On applying them to musical temporal spaces, one has to make certain concessions by removing all possible asymmetrical directional properties. One aspect of doing this, is to consider mainly absolute-values of event properties and intervals, for example, absolute-value pitch intervals (i.e. not direction: ascending-descending). Of course, event start-times must be omitted as these are the basic directional temporal values, and absolute start-time intervals should be used instead.

There is though one aspect of musical asymmetry that cannot be avoided. This relates to the fact that musical objects are asymmetric object themselves (even the most simplified homogeneous description of a note distinguishes between its attack and the rest of its body). This asymmetry is reflected in that, for example, the temporal grouping rules can never give an identical grouping structure to the original and the retrograde form of a melody - unless one considers a melody of 'timeless' musical objects! It relates to the way that rules of perceptual organisation give different grouping boundaries for musical duration sequences and for start-time interval sequences. It will be shown below how the interaction between these different groupings results in the asymmetric perceptual organisation of a sequence of musical events.

We will now attempt to define the Identity-Change rule (lower-level and more precise than the Gestalt principles) and the Proximity rule (similar to the Gestalt principle of proximity extended for any kind of interval) which will form the basis of the *LBDM*. These rules will be discussed initially for any sequence of two or three objects and then will be applied to sequences of musical objects.

3. The *Local Boundary Detection Model*

3.1 The *Identity-Change* and *Maximum Interval Rules*

Let us consider two discrete symmetrical objects, and an external observer who has selected a set of properties that are pertinent to the objects' description and a set of metric units with which discrete quantised (perceptually pertinent) values for each of these properties can be calculated. If the values for all the properties are the same (*except* for the spatio-temporal values/coordinates) then the two objects are said to be *identical.* If at least one is different then they are *non-identical.*

It is assumed that our observer is only concerned with attributes and relations that are immanent to the two-object world. Obviously spatial or temporal coordinate values rely on the observer's system of reference and thus are disregarded. Absolute values, though, of spatial or temporal *intervals* (distances) between two objects may be considered as they are independent of any external reference axes. Absolute values of intervals relate to what was discussed in the previous section.

In the case where the two objects are identical (fig. 2a) the observer is not able to tell them apart and to label them solely on their immanent attributes. This is a consequence of the fact that they both have all properties the same i.e. there isn't a single property which is unique to either of them. (We do not consider properties that may arise in relation to the observer's viewpoint - e.g. left-right, up-down, near-far). If the objects are non-identical (fig. 2b) then the observer may give them different descriptions and label them uniquely. In this case a *change* is detectable/discernible and a possibility arises for placing the two objects in different classes/groups.

figure 2

a. • •   *b.* • +

c. • • •   *d.* • •   •   *e.* • +   +

Let us suppose now that a third object is added to the two-object universe and that it's placed on the same axis with the other two (fig. 2c, d, e). The observer can now not only compare the property values of each pair of adjacent objects but the intervals that are defined by each pair as well i.e. inferences can be made as to whether two intervals are identical or not. And what was previously discussed about the two objects can now be applied to two intervals (including space/time intervals).

For the three identical objects of *fig. 2c* the observer may calculate that the two spatial/temporal intervals are identical and consequently that there is no way to group the three identical objects in uniquely defined sub-groups. On the contrary, in *fig.2d* the two distances are non-identical - i.e. a change is discernible between them - and this change makes it possible to uniquely define and identify any sub-group of adjacent objects relying solely on the immanent features of the three-object world (e.g. find the group that consists of the two adjacent objects that are closer together).

Assuming that we are presented with any sequence of values that are relating to the values of properties of objects or intervals between such values, grouping boundaries may be inserted in the sequence according to the following rule:

> ***Identity-Change Rule****:* Grouping boundaries may be introduced only between two non-identical elements (objects, parametric values, intervals). Identical elements do not allow any such boundaries between them.

The traditional Gestalt rules of proximity and similarity are higher-level than the above rule as they can be applied on *at least* three objects. The main claim made here is that facets of the similarity principle are captured by the ICR rule when this is applied to consecutive pairs of objects or intervals (more on this issue in the next section).

The Identity-Change rule is partially supported by an experiment [Garner 1974, quoted in Handel 1989] wherein an eight-element pattern composed of two different pitch elements, for example XXXOXOOO, is looped indefinitely and listeners are asked to describe the pattern they perceive. Various preferential ways of organisation were recorded (there are eight possibilities starting on each element of the sequence) but hardly ever did any listener break a run of same elements.

When the application of ICR on two consecutive intervals detects a change and suggests a local boundary, this boundary is ambiguous i.e. the middle object can be placed on either side of the boundary (for example, in *fig. 2d* ICR suggests a boundary on the first and/or the second interval). How can a decision be taken as to whether the grouping boundary should occur on the one or the other side of the object? This is where the second fundamental rule (relating to the Gestalt rule of proximity - and similarity - seen as local maxima in interval sizes) comes into play:

> ***Proximity Rule****:* Amongst three adjacent objects those two will tend to form a group that are closer together (or more similar to eachother) i.e. a boundary will be inserted on the larger interval.

In *fig. 1d* the two objects on the left will tend to be grouped together (i.e. the PR rule gives a grouping boundary between the two objects on the right).

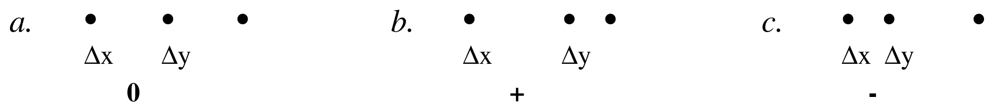
3.2 Applying the ICR and PR rules on three-event musical sequences.

We will assume that for each parametric feature of a musical surface we can construct a sequence of intervals on which the ICR and PR rules may be applied. We will start off by presenting the application of the rules to the following intervals defined by parametric properties of contiguous musical events: pitch, dynamics, rests and articulation. The grouping boundaries resulting from the sequence of start-time intervals and durations will be presented at the end of this section.

The pitch intervals may be measured in semitones or scale-steps depending on the kind of melody at hand (tonal, atonal etc.). We use the absolute values, i.e. the interval size regardless of direction (ascending-descending). For the loudness intervals, we use the scale from *ppp*=1 to *fff*=8 to calculate the absolute value of the interval between these scale values (e.g. *mp-ff* is 3 units). Rests are considered as the distances between the end-time of the previous and the start-time of the next event (the maximum common denominator of the time intervals appearing in a musical surface can be used as a unit e.g. 16th durational value = 1). Slurs, staccati, breath-marks etc. are considered to be expressional rests and are inserted between the notes they mark as normal rests that have a value that is a fraction of the preceding note (we will not introduce here any detailed way of measuring them).

The relation between two intervals can be of two types: *identity* or *change.* For reasons of asymmetry that will be introduced later on we will depict the change relation in two directional forms (fig. 3b,c). In the following figures, dots represent parametric values of musical events and the distances between the dots the interval sizes between these values (Δx, Δy are interval values and are placed at the left-hand side of the interval). In *fig. 3a* Δx=Δy and the identity relation is represented by a zero. In *fig. 3b* Δx>Δy and in *fig. 3c* Δx<Δy, and the change relations are represented by the + and - signs respectively.

figure 3



	a.		b.		c.	
ICR:	0	0	1	1	1	1
PR:	0	0	1	0	0	1
Total	0	0	2	1	1	2

At this stage we will introduce numeric values for the strength of the ICR and PR rules (more research is necessary for the selection of the most appropriate values). A numeric value is given to each interval as shown below:

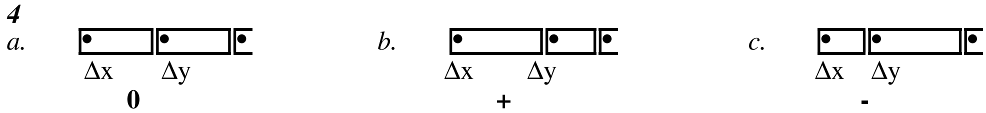
ICR: 0 for the identity relation (0 for each interval)
2 for the change relations (1 for each interval)

PR: 0 for no boundary
1 for proximity boundary suggestion

We get thus the total interval boundary preference strengths as depicted in *fig. 3* (last three lines).

We can now examine the duration and start-time intervals. The duration of a musical event is an internal attribute of that event whereas start-time intervals are temporal distances between two different successive events (a rest is considered to be part of the preceding note's duration - rests are considered independently in a separate parametric profile). We have thus the application of the ICR and PR rules for the start-time intervals exactly as described above plus the application of ICR for the sequence of durations (numeric strength 2). We now have the following kinds of relations for two start-time intervals delimited by 3 start-time points (dots) and the two corresponding durations (rectangles) (fig. 4).

figure 4



	a.		b.		c.	
ICR (st-ints)	0	0	1	1	1	1
PR (st-ints)	0	0	1	0	0	1
ICR (dur)	0	0	2	0	2	0
Total	0	0	4	1	3	2

It is now clear that the + and - change relations are not symmetric. It is not possible to apply the principles of perceptual organisation in the musical domain without introducing local asymmetry.

3.3 Applying the ICR and PR rules on longer musical surfaces

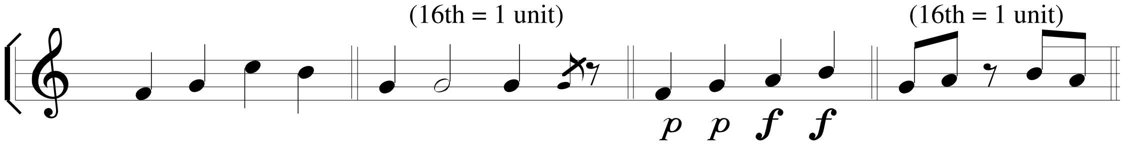
For a given parametric profile of a musical surface one finds all the kinds of interval relations (0, +, -) that exist between every two successive intervals. If there are 3 or more consecutive + or - (e.g. +++, - - - - -), then only the ones at the ends are considered - the others do not contribute to the numeric strengths. Then, the sum of the numeric strengths for each kind of relation is calculated. For a single numeric strength sequence *all* the local maxima suggest the most preferable local grouping boundaries. If a local maximum consists of two numbers, then this is an ambiguous boundary (explained in the next section). If a single maximum has an adjacent numeric strength that differs by one unit, then this suggests a possible ambiguous boundary. If three numbers constitute a local maximum or one maximum is surrounded by two numeric values that differ by one numeric unit then the middle interval is preferred as a grouping boundary (it is not possible to have a maximum with more than three numbers).

The above procedure is realised for every parametric profile of interest. Then the *total* sum of all the numeric strength strings is calculated. The local peaks are the points in a melodic sequence in which boundaries may preferably appear. Obviously the *LBDM* rules may contradict each other for different parametric profiles leading to local peaks consisting of more than one similar values. In this case one may estimate the relative strength of parametric profile and decide which wins or simply consider this boundary as ambiguous. As in this study we do not aim at reaching a final segmentation of a given surface (higher level grouping rules have to be taken into consideration) we will simply leave ambiguities unresolved.

The 0, +, - interval relations can be combined in pairs to form sequences of three intervals (i.e. 4 events). There are 3^2=9 possible combinations. In *fig 5* we give a first example of how one can use the ICR & PR rules to calculate the strengths of grouping boundaries for the <u>- +</u> combination.

figure 5

a. scale-step ints | b. start-time ints (16th = 1 unit) | c. dynamic ints | d. rest intervals (16th = 1 unit)

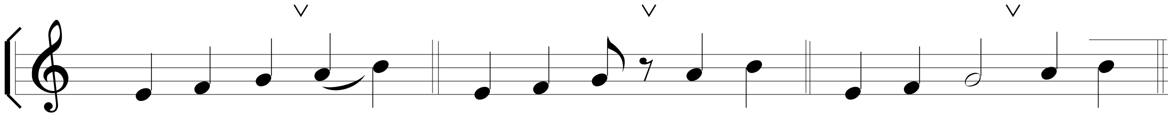


	a	b	c	d
intervals	1 3 1	4 8 4	0 3 0	0 2 0
	- +	- +	- +	- +
'-' values	1 2	3 2	1 2	1 2
'+' values	2 1	4 1	2 1	2 1
sum:	1 **4** 1	3 **6** 1	1 **4** 1	1 **4** 1
	^	^	^	^

As it happens, almost all the grouping preference rules of Lerdahl & Jackendoff and all the grouping rules suggested by Tenney & Polansky fall under the <u>- +</u> category! (Exception: L&J's GPR3d (equal note length) and the articulation changes form legato to staccato and the opposite, fall under the 0 + 0 and 0 - 0 combinations). See *fig. 6* for the application of the *LBDM* rules to the local detail examples of Lerdahl & Jackendoff's grouping theory. *LBDM* accounts for all the positive instances of the Lerdahl & Jackendoff's groupings and for all the examples where their rules do not apply!

figure 6: Application of the *Local Boundary Detection Model* to the Lerdahl &Jackendoff [1983, p.44-46] local detail grouping examples 3.14-3.17. For the examples not accounted for by the GPR2 and GPR3 rules, the proposed theory suggests ambiguous boundaries (depicted as ^__^).

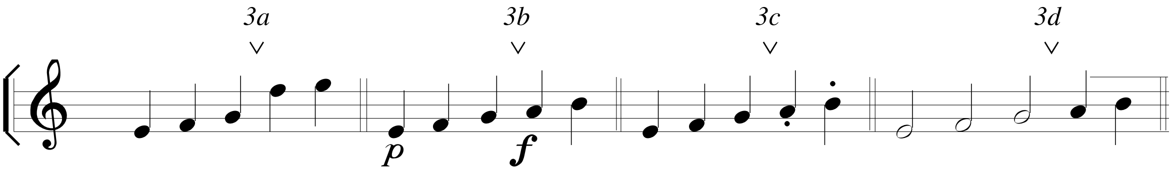
L&J rules:	*2a* v	*2a* v	*2b* v



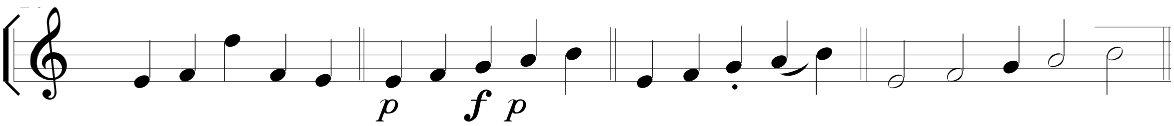
	0 0 1 0	0 0 2 0	4 4 8 4 4
	0 - +	0 - +	0 - + 0
LBDM:	0 1 **4** 1	0 1 **4** 1	0 3 **6** 1
	^	^	^

(3d)

0 1 1 0	0 2 2 0	4 8 8 4 4
- 0 +	- 0 +	- 0 + 0
1 **2** **2** 1	1 **2** **2** 1	3 2 **4** 1
^___^	^_____^	^



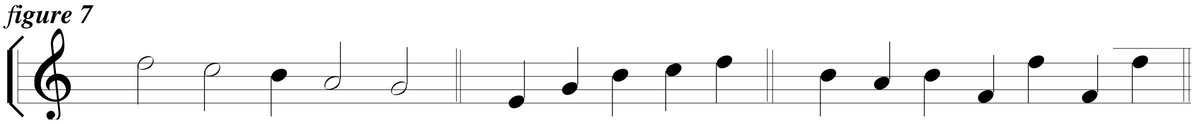
1 1 4 1	0 0 3 0	0 0 1 3 3	8 8 8 4 4
0 - +	0 - +	0 - - 0	0 0 + 0
0 1 **4** 1	0 1 **4** 1	0 1 **3** **2**	0 0 **4** 1
^	^	^__	^



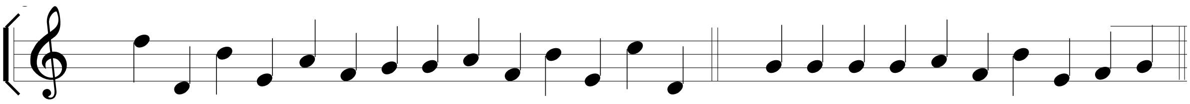
1 5 5 1	0 3 3 0	0 1 3 0	8 8 4 8 8
- 0 +	- 0 +	- - +	0 + - 0
1 **2** **2** 1	1 **2** **2** 1	1 **3** **4** 1	0 **4** **4** 2
^__^	^__^	__^	^____^

The - + interval relation is the strongest no matter what the interval context in which it appears is. The rest though of the possible combinations are important as well, especially when they are enclosed between two 0's (identical intervals). Most of the existing grouping theories do not account for these cases. In *fig. 7* *LBDM* boundaries are depicted for the following sequences: 0 + - 0, 0 + 0, 0 - 0, 0 + + 0, 0 - - 0 (runs of + and - that do not contribute to the strength values are enclosed in parentheses).

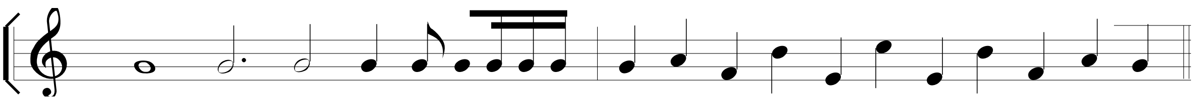
figure 7



8 8 4 8 8	2 2 1 1	1 1 3 5 5 5
0 + - 0	0 + 0	0 - - 0 0
0 **4** **4** 2	0 **2** **1**	0 1 **3** **2** 0
^___^	^__	^__



6 5 4 3 2 1 0 1 2 3 4 5 6	0 0 0 1 2 3 4 1 1
(+ + + + +) + - (- - - - -)	0 0 - (- -) - + 0
0 0 0 0 0 2 **2** 2 0 0 0 0	0 0 **1** **2** 0 1 **4** 1
^	__^ ^

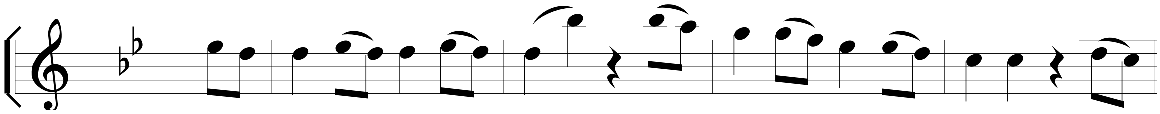


16 12 8 4 2 1 1 1 1	1 2 3 4 5 5 4 3 2 1
(+ + + +) + 0 0 0	(- - -) - 0 + (+ + +)
0 0 0 0 **4** 1 0 0	0 0 0 1 **2** **2** 1 0 0
^	^__^

In the following examples the *LBDM* is applied on musical excerpts from classical and contemporary pieces. The preferred grouping structure is presented for Mozart's opening of the *Symphony in G min.* (fig. 8), an excerpt from Xenakis' *Keren* (fig. 9) and an excerpt from Stravinsky's *Three pieces for solo clarinet,* no. III (fig. 10).

figure 8: Low-level grouping structure for the theme of Mozart's *Symphony in G min.* (sequence A). Sequence B depicts the most probable local boundaries given by the refined version of *LBDM* (see last section of this paper for further details).

								3d		2a								3c		2a	
L&J rules:			2b			2b		2a	3a	2b			2b			2b		3d		2b	
			v			v		v	v	v			v			v		v		v	



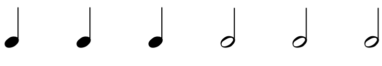
start-time ints:	0	-	+	0	-	+	0	-	-	+	0	-	+	0	-	+	0	-	-	+	
	0	3	6	1	3	6	1	3	5	6	1	3	6	1	3	6	1	3	5	6	(1)
scale-step ints:	+	-	0	+	-	0	+	-	+	-	0	+	-	0	+	-	0	+	-	0	
	2	2	2	2	2	2	2	2	4	2	2	2	2	2	2	2	2	2	2	2	
rests (slurs):	+	0	+	-	0	+	-	+	-	+	-	0	+	-	0	+	-	+	-	+	
	1	2	2	2	2	2	2	4	2	4	2	2	2	2	2	2	2	4	2	4	(1)
Total: A.	3	7	**10**	5	7	**10**	5	9	11	**12**	5	5	**10**	5	7	**10**	5	9	9	**12**	
			^			^				^			^			^				^	
B.	.4	1.7	**2.8**	.5	1.7	**2.8**	.5	2.3	2.6	**5.7**	.6	1.7	**2.8**	.5	1.7	**2.8**	.5	1.8	2.5	**4.9**	
			2.8			**2.8**				5.7			2.8			2.8				4.9	
										5.7										4.9	

3.4 Further comments on the application of the *LBDM* rules

• Most grouping formal theories define exclusively clear boundaries that appear unambiguously between two musical objects. There are cases though where a boundary is ambiguously suggested. This phenomenon is conveniently accommodated within the present theory wherein numeric peaks with two identical or similar values suggest a blurred boundary (higher level grouping mechanisms may support one interpretation over other possibilities). [Deliège 1987, p. 331, 342] suggests that in the following sequences (fig. 11) the grouping boundary perceived by listeners tends to appear after the first half-note and staccato note respectively. The current theory suggests an ambiguous boundary on those notes.

figure 11

L&J rules: 3d



0	0	-	0	0
0	0	**3**	**2**	0
		^____		

3c

0	0	-	0	0
0	0	**1**	**2**	0
			____^	

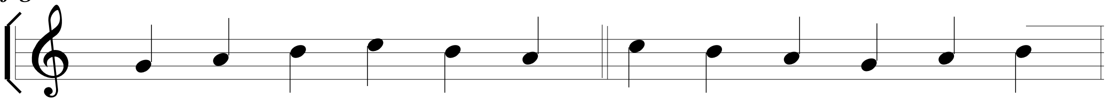
• It may be preferable in some cases to use subjective scales for interval sizes instead of acoustic ones. For example, in the following series of equally timed elements (fig. 12) the more intense ones tend to be perceived as beginnings of groups [Handel, 1989, p. 386-389]. In other words, it may be said say that the interval *p* —> *f* is larger than the reverse *f* —> *p*. The sequence below will have the following grouping boundaries:

figure 12

	f		*p*		***f***		*p*		***f***		*p*		***f***		*p*	
physical intervals	3		3		3		3		3		3		3			
		0		0		0		0		0		0				(no boundaries)
	f		*p*		***f***		*p*		***f***		*p*		***f***		*p*	
subjective intervals	2		3		2		3		2		3		2			
		-		+		-		+		-		+				
		1		**4**		1		**4**		1		**4**		1		
				^				^				^				

• [Deliège 1987] suggests that a change in melodic contour contributes weakly towards the establishment of a local boundary. This may be incorporated in the current theory by detecting changes of contour of the form 0 * 0, e.g. U U D D, and at the point of change applying the ICR rule - 1 numeric value for each interval (fig. 13a).

figure 13



	U	U	U	D	D		D	D	D	U	U
	0	0	*	0			0	0	*	0	
a)	0	0	1	1			0	0	1	1	
			^	^					^	^	
b)	0	0	2	1			0	0	2	1	
			^						^		

Deliège [1987, p.353] reports that the analysis of the responses of listeners to the change of the melodic contour "revealed a preference for cutting *before* the pivot sound". Taking this observation into account it would seem plausible to give an extra numeric weight at the first interval (fig. 13b).

4. Refinements in the *Local Boundary Detection Model*

The *LBDM* can be enhanced in various ways so as to accommodate further nuances of musical perception that contribute towards a more accurate description of the low-level grouping structure of a musical surface. Some of these are described below:

1. The various parametric profiles may be given different weights depending on the degree of prominence they may have for a given melodic surface. If, for instance, start-time intervals are considered more important, then the start-time profile may be given a higher weight factor before it is added to the other profiles.
2. The numeric value of the PR rule may be augmented (e.g. have a value of 2). This will produce sharper local maxima.
3. The 0, +, - identity/change relations may be refined by taking into account the ratio/difference between two interval sizes (factor α). As [Deliège 1987, p. 328] points out, the sensation of a boundary is strengthened in correspondence to the increase in difference between two intervals. For example, the second of the following two sequences suggests a stronger boundary:

4. A further factor that contributes to the perceived strength of a boundary relates to the total sum of the two intervals. The larger the sum is, the greater the prominence of the perceived boundary (factor β). For example, the second of the following two sequences suggests a stronger boundary:

5. Values that appear under slurs may be attenuated or deleted altogether.

Taking in account some of the above suggestions (mainly factors α & β) the local boundary strengths have been calculated for fig. 8 & 10 (if x, y are positive integer interval sizes then factors α & β may be calculated using functions such as: $\alpha = (x-y)/(x+y)$ and $\beta = 1 - 1/(x+y)$ and $0<\alpha,\beta<1$). For the theme of Mozart's *Gmin Symphony* it is clear that the middle and last boundaries are more prominent and could be considered as best candidates for higher level groupings (actually, these boundaries would emerge if the second-order local maxima were selected i.e. the maxima of the first-order maxima). This is a rather interesting result, especially if one bears in mind that no higher level perceptual principles have been employed (e.g. symmetry, parallelism).

A second example is given for an excerpt from the 3rd piece from the *Three pieces for solo clarinet* by I. Stravinsky. Lerdahl and Jackendoff apply their grouping preference rules on the beginning of the 1st of these pieces to show that their theory is general and not style specific. If though a different excerpt from this set of monophonic pieces (fig. 10) is examined the local boundaries proposed by Lerdahl and Jackendoff show limitations in two respects: firstly, not all the perceptually significant points of segmentations are accounted for (see, for example, the third grouping boundary - after the 10th note); secondly, many points are given excessive grouping boundary importance (see, for example, the second half of the excerpt in which strong GPR 2a and 2b boundaries are placed on every rest). On the contrary, the refined version of *LBDM* accounts for all the possible local boundaries and also highlights those that are more prominent (actually, as it happens, the second-order maxima suggest boundaries which correspond to the composer's articulations).

The refined *LBDM* encompasses facets of similarity more effectively as it accounts for the degree of difference between two intervals. The refined *LBDM* may be incorporated in real-time systems that attempt to segment input musical data. If, for instance, two input durations are almost the same - but not identical - factor α will tend to become zero so this slight performance difference will not contribute towards the establishment of a boundary (there is no need for quantisation of musical parameters before segmentation). It can also cope with the longer strings of only + or - change relations (e.g. ++++) because these changes will receive different strengths according to their relative factor importance. In addition to musical sequences, it seems plausible that the proposed model may be applied to any sequence of visual or auditory events.

5. Conclusion

In this paper a formal theory has been described that attempts to define local boundaries in a given melodic surface. The proposed *Local Boundary Detection Model* is based on two rules, namely the Identity-Change and the Proximity rules, and it is suggested that it presents (especially the refined version) a more effective and general method for low-level segmentation in comparison to other existing models.

References

[Bregman 1990] A. S. Bregman, *Auditory Scene Analysis*. MIT Press, Cambridge (Ma], 1990.

[Deliège 1987] I. Deliège, Grouping Conditions in Listening to Music: An Approach to Lerdahl and Jackendoff's Grouping Preference rules. *Music Perception,* 4, 325-360

[Deutsch 1982a] D. Deutsch, Grouping Mechanisms in Music. In D. Deutsch (Ed.), *The Psychology of Music,* Academic Press, New York, 1982.

[Deutsch 1982b] D. Deutsch, The Processing of Pitch Combinations. In D. Deutsch (Ed.), *The Psychology of Music,* Academic Press, New York, 1982.

[Deutsch & Feroe 1981] D. Deutsch & J. Feroe, The Internal Representation of Pitch Sequences in Tonal Music, *Psychological Review,* 88, 503-22.

[Dowling & Harwood 1986] W. J. Dowling & D.L. Harwood, *Music Cognition.* Academic Press, Orlando (Fl), 1986.

[Garner, 1974] W. R. Garner, *The processing of Information and Structure,* Erlbaum, Potomac (Md), 1974.

[Handel 1989] S. Handel, *Listening. An introduction to the perception of auditory events.* MIT Press, Cambridge (Ma), 1989.

[Lerdahl & Jackendoff 1983] F. Lerdahl & R. Jackendoff, *A Generative Theory of Tonal Music.* MIT Press, Cambridge (Ma), 1983.

[Lewin 1987] D. Lewin, *Generalized Musical Intervals and Transformations.* Yale University Press, New Haven, 1987.

[McAdams 1984] S. McAdams, The auditory Image: A metaphor for musical and psychological research on auditory organisation. In Crozier, W. R. & Chapman, A.J. (Eds), *Cognitive Processes in the Perception of Art.* North-Holland Press, Amsterdam, 1984.

[Tenney 1961] J. Tenney, *Meta+Hodos.* Frog Peak Music, Oakland (Ca), 1964.

[Tenney & Polansky 1980] J. Tenney & L. Polansky, Temporal Gestalt Perception In Music, *Journal of Music Theory,* 24, 205-241.

[Xenakis 1992] I. Xenakis, *Formalised Music,*.Pendragon Press, New York, 1992.