Music Perception
Winter 2001, Vol. 19, No. 2, 169–198

Perceptual Mechanisms in Music Processing

DIRK-JAN POVEL & ERIK JANSEN
University of Nijmegen, The Netherlands

To gain a better understanding of the processes by which human listeners construct musical percepts within the Western tonal system, we conducted two experiments in which the perception of brief tone series was studied. The tone series consisted of (fragments from) different orderings of the collection C_4 E_4 $F\sharp_4$ G_4 $B\flat_4$ and were preceded by two chords to induce a key. Two different tasks were used: (1) rating the melodic "goodness" of the tone series and (2) playing a few tones that complete the tone series. In Experiment 1, tone series of different lengths were presented in blocks. In Experiments 2a and 2b, increasing fragments of tone series were presented to examine the development of musical percepts. The majority of the data can be explained by two perceptual mechanisms: *chord recognition* and *anchoring*. Chord recognition is the mechanism that describes a series of tones in terms of a chord, a mental unit stored in long-term memory. Anchoring is the mechanism by which a tone is linked to a tone occurring later in the series. The paradigm appears to be a powerful tool for tracing perceptual mechanisms at work in the on-line processing of music.

Received January 4, 1999, accepted December 13, 2000.

A principal goal of research in music perception is to reveal the mental representations that listeners form of musical pieces and to discover the perceptual processes that yield these representations. Two important representational theories of music perception, that of Lerdahl and Jackendoff (1983) and that of Deutsch and Feroe (1981), assume that the listener's mental representations are based on Gestalt laws, on the one hand, and on various structural regularities present at different hierarchical levels in the music, on the other hand. Research in music perception has indeed shown that Gestalt laws play an important role in the processing of music (Deutsch,

Address correspondence to D. J. Povel, Nijmegen Institute for Cognition and Information (NICI), P. O. Box 9104, 6500 HE Nijmegen, The Netherlands. (e-mail: povel@nici.kun.nl)

ISSN: 0730-7829. Send requests for permission to reprint to Rights and Permissions, University of California Press, 2000 Center St., Ste. 303, Berkeley, CA 94704-1223.

1982), as do various types of regularities in musical structure (e.g., Bigand, 1990; Cuddy, Cohen, & Mewhort, 1981; Deliège, 1992, 1996; Deutsch, 1980; Halpern, 1984; Povel & Van Egmond, 1993; Serafine, Glassman, & Overbeeke, 1989; Sloboda & Parker, 1985; Tan, Aiello, & Bever, 1981; see also Howell, West, & Cross, 1991).

Although the theories of Lerdahl and Jackendoff (1983) and Deutsch and Feroe (1981) can, in principle, be used to describe the mental representations that music listeners form, they cannot be used to predict these representations. That is, if one knows how a piece of music is perceived, the theories can render an appropriate description for that interpretation, but they cannot indicate that interpretation in advance. In fact, instead of generating a unique representation, both theories yield multiple representations for one piece of music. In the theory of Lerdahl and Jackendoff, these multiple interpretations result from the fact that several different rules may apply to a tone sequence and that the rules interact in various ways. Likewise, in Deutsch and Feroe's coding language, in which a piece is described as operations performed on alphabets of tones, different representations for a single melody may be obtained, depending on the choice of alphabets, operations, and reference points. The sequence B_4 C_5 $D\sharp_5$ E_5 $F\sharp_5$ G_5, for instance, could be described as a series of tones in the chromatic scale in which the sizes of the different steps are specified, but also as an arpeggiated C chord (C_5 E_5 G_5) in which the tones B_4, $D\sharp_5$, and $F\sharp_5$ are ornaments of the three chord tones. The latter interpretation is what listeners will usually report hearing (Laden, 1994). Moreover, Deutsch and Feroe's coding language provides a code (mental representation) for every tone sequence, irrespective of it being perceived as a musical percept.

One motive for the present study has been the inability of current theories of music perception to generate precise predictions. What is needed is a better insight into how listeners process music. By closely studying listeners' reactions to tone sequences of increasing size, we want to discover the perceptual mechanisms central to the processing of music. The insights obtained in this way may eventually lead to more specific theories of music perception.

THE PROCESSING OF MUSIC

Before dealing with the study proper, we briefly discuss a few fundamental aspects of music perception. Music perception may be defined as the process that aims to transform a series of unconnected tones into an integrated mental representation in musical terms. A sequence of sounds that is conceived musically (rather than linguistically or otherwise) will be mapped on two dimensions: the *pitch-height dimension* (yielding the pitch of the sound) and the *key dimension* (yielding the attribute of scale degree). The key dimension is the hierarchically organized mental tone space in which the relations between tones and chords are specified (Krumhansl, 1990). It serves as an interpretational frame that supplies the musical function of the sounds. As soon as the pitches of a sequence have activated a specific key, those pitches are identified as tones in a scale. All tones in a key are associated with a certain degree of "stability" (e.g., the first tone of the scale is the most stable tone, and the last tone, the "leading tone," is the least stable) and with a tendency to resolve to other tones (Cooke, 1959; Povel, 1996; Zuckerkandl, 1956). Thus, in making a musical interpretation of a tone series, the tones function simultaneously in these two dimensions. Each dimension plays a role in the formation of musical percepts.

The main characteristic of the pitch-height dimension that contributes to melody formation is obviously pitch, especially pitch proximity. Melodies tend to proceed by step, rather than by leap, thus forming smooth contours; segmentation of tones is based in part on the principle of proximity: tones relatively close in pitch will form perceptual clusters or groups.

The characteristics of the key dimension that contribute to melody formation are related to the properties of scales and chords. The availability of a scale enables one to describe a sequence of consecutive scale tones as a "run" using the next relations between the tones (the sequence C_5 D_5 E_5 F_5 G_5, for instance, can be represented as $4N(C_5)$: start with C_5 and add four times the next (N) element). The key dimension also allows the description of (a part of) a series of tones as a chord. Thus the first three tones of the series C_5 E_5 G_5 A_5 B_5 C_6 may be recognized as a major triad and encoded as such. Also other types of chords may play a role in the encoding of tone series. Another principle associated with the key dimension is that of anchoring. This mechanism, first described by Bharucha (1984), links or "anchors" an unstable tone to a more stable tone. Anchoring is based on the notion of a hierarchical relationship between the tones in a key, with, for instance, the diatonic tones (the tones of the scale) being hierarchically higher than the chromatic (nonscalar) tones (Lerdahl, 1988). Tones lower in the hierarchy, the less stable tones, are attracted by more stable tones higher in the hierarchy (Povel, 1996; Zuckerkandl, 1956). Bharucha (1984, 1996) has shown that a tone can be anchored only to a tone very close in pitch (1 or 2 semitones away) that follows the tone to be anchored, usually (but not necessarily) the immediately succeeding tone. Phenomenally, the anchored tone is perceived as an ornament of the tone to which it is anchored.

The process of music perception can now be conceived as follows: Given a tone sequence presented in a specific tonal context, a search is initiated for perceptual mechanisms that establish relations between the tones. If all tones of the sequence can be related to each other in this way, an integrated musical percept results.

THE EXPERIMENTAL PARADIGM: GOODNESS JUDGMENTS AND EXPECTATIONS

To study the processes involved in tonal music perception, we used two different tasks to examine the perception of tone sequences of different lengths. The first task required subjects to judge the "goodness" of a tone sequence (Experiments 1 and 2a). The second task was to play a few tones that completed the sequence (Experiment 2b). To study more closely the on-line percept formation, in Experiments 2a and 2b we used a "gating" paradigm. In this paradigm, a tone sequence is presented in stepwise increasing fragments, and responses are collected for each successive fragment.

PRELIMINARY STUDY

In an exploratory study (Povel, 1997), 12 subjects were presented with 53 of the 120 permutations of the sequence C_4 E_4 $F\sharp_4$ G_4 $B\flat_4$ (the subscript 4 denotes that the tones are in the fourth octave beginning with C_4 of 261.6 Hz). By using the gating method, each tone series was presented in fragments of increasing length, beginning with length 3, and each fragment was preceded by the chords C7 and F to induce the key of F major. In the trials, the stimuli were presented at different pitch heights. For each fragment, the subjects indicated whether or not the series led to a musical percept (Figure 1).

An analysis of the fragments that were judged as being musical percepts indicated that two types of tone series led to musical percepts, suggesting two perceptual mechanisms. The first type of fragment contained only the tones C E G or the tones C E G B♭, in any order (the results are here interpreted as if all series were presented in the key of F major). This suggested *chord recognition* as the underlying perceptual mechanism: If a series is

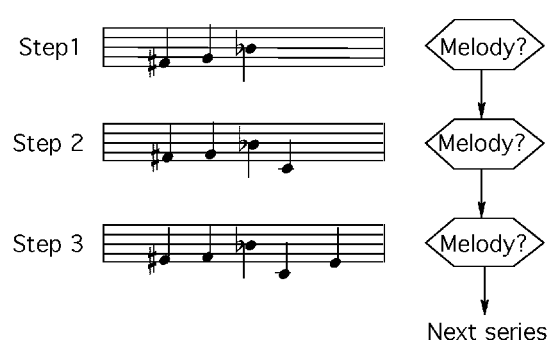


Fig. 1. The gating paradigm used in the pilot experiment and in Experiments 2a and 2b. ("Melody?" means: does the fragment lead to a musical percept?)

recognized as either a triad (C E G) or a dominant seventh chord, (C E G B♭), it may be coded as a perceptual chunk and thus result in a musical percept. The second type of fragment contained the tone F♯ (which does not belong to the scale of F major) followed by a G. This suggested the mechanism of *anchoring* discussed earlier, whereby the F♯ is linked to the G. The operation of the mechanisms of chord recognition and anchoring in the perception of the tone series F♯ G C B♭ is schematically displayed in Figure 2.

PRESENT INVESTIGATION

In the experiments reported next, we have examined to what extent the two hypothetical perceptual mechanisms, chord recognition and anchoring, can explain the perception of a set of tone series. In addition, we wanted to obtain a detailed insight into the mental representations and the expectations that listeners develop while listening to tone series.

Experiment 1

The main purpose of this experiment was to determine to what extent chord recognition and anchoring can explain the perception of brief tone series derived from the collection C_4 E_4 $F\sharp_4$ G_4 $B\flat_4$ presented in the context of F major. In particular, we tested the following hypotheses: (1) fragments containing only chord tones will be rated highest because the mechanism of chord recognition can be applied; (2) fragments containing an F♯ followed by a G (either immediately or with intermediate tones) will be rated lower, because two mechanisms (chord recognition and anchoring) must be applied.[1]

Fragment size	Input	Process(es)	Percept
3	F♯ G C	Anchoring	F♯ ↘ G C
		Chord recognition	{G C} → fifth or C triad
4	F♯ G C B♭	Chord recognition	{G C B♭} → C7

Fig. 2. Hypothetical operation of the perceptual mechanisms *chord recognition* and *anchoring* in the incremental processing of the tone series F♯ G C B♭. Upon hearing the fragment F♯ G C, the F♯ is anchored to G and G C is recognized as a chord. When subsequently the B♭ is added, the tones G C and B♭ are recognized as C7.

1. The rationale for this hypothesis is that more perceptual effort is needed if two mechanisms must be applied rather than just one: these two perceptual mechanisms must be activated by specific features of the stimulus. The chord recognition mechanism must be put on hold when a nonchord tone occurs. Thus the listener must try harder to "understand" the sequence, which is reflected in a lower goodness rating.

Further it is hypothesized that listeners' ratings will decrease with the number of intermediate tones; (3) fragments containing an F♯ not followed by a G will be rated lowest, as neither the perceptual mechanism of chord recognition nor that of anchoring can be applied.

METHOD

Participants

Twenty subjects, graduate and undergraduate students at the University of Nijmegen, participated in the experiment. The median age of the subjects was 23.5 years. None of them were professional musicians, and none had received any formal music-theoretical training. All of them had played a musical instrument for several years, with a mean of 10.2 years.

Stimulus Materials

Stimuli were selected from the 120 permutations of the tone set C_4 E_4 $F♯_4$ G_4 $B♭_4$. This set of tones (which was also used in the pilot study) was chosen because it seemed most suited for studying the on-line processes that occur in music perception. First of all, the tones in the set do not belong to any one major scale (C E G and B♭ are all members of the scale of F major, but F♯ is not), although they do all belong to the scale of G ascending melodic minor. Second, different subsets of tones within the set may evoke the two dominant seventh chords C7 and F♯7 (the tones C E G and B♭ are elements of the C7 chord, whereas the tones E A♯ [B♭] and F♯ are elements of the F♯7 chord) that are the dominants of F and B, respectively, which are remote keys located on opposite sides of the circle of fifths. In consequence, the sequence is musically quite ambiguous, as a result of which percept formation will probably depend highly on the order of the tones in the sequence as well as on the induced key.

Stimulus Selection

To begin with, all initial fragments of length 3, 4, and 5 of all permutations of the combination C_4 E_4 $F♯_4$ G_4 $B♭_4$ were assigned to the following 7 categories:

1. Chord recognition 1 (Ch1): the fragment consists of some permutation of the three notes of the triad C E G. This category contains only fragments of length 3.
2. Chord recognition 2 (Ch2): the fragment consists of some combination of notes from the C7 triad, provided it includes B♭ (otherwise it would fall in category 1). This category contains fragments of length 3 and 4 only.
3. Anchoring 1 (A1): the fragment contains the consecutive tones F♯ and G. This category contains fragments of all three lengths.
4. Anchoring 2 (A2): the fragment comprises the tone F♯, which after one intermediate tone is followed by the tone G. This category contains fragments of all three lengths.
5. Anchoring 3 (A3): the fragment comprises the tone F♯, which after two intermediate tones is followed by the tone G. This category contains fragments of lengths 4 and 5.
6. Anchoring 4 (A4): the fragment comprises the tone F♯, which after three intermediate tones is followed by the tone G. This category contains only fragments of length 5.
7. No melody (N): all fragments not belonging to one of the first six categories. This category contains fragments of all three lengths.

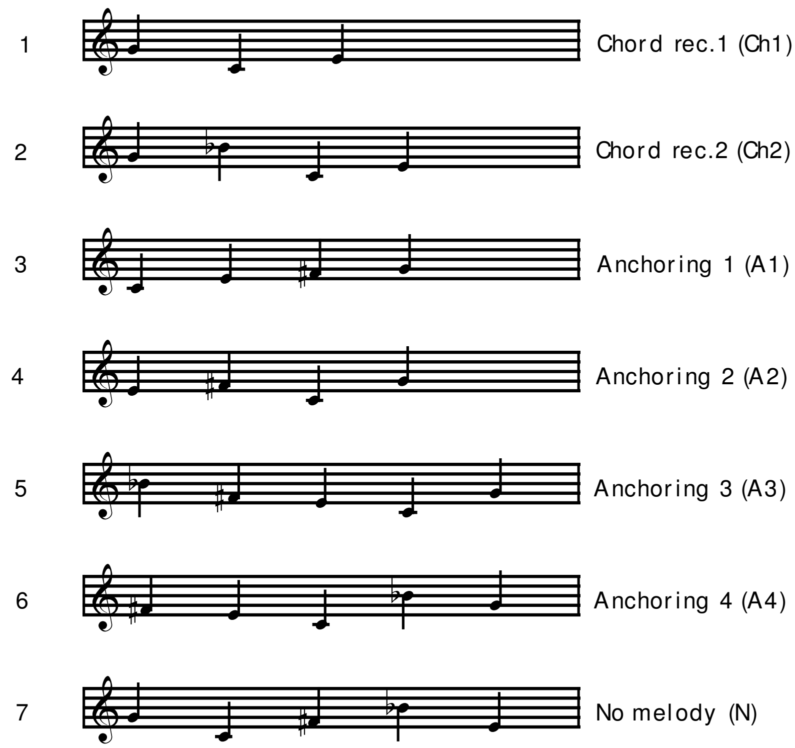


Fig. 3. Examples of tone series in each of the 7 different categories distinguished in Experiment 1.

Sample sequences for each of the seven categories are shown in Figure 3.

Next, for each of the fragment sizes (3, 4, and 5), we selected 5 series from each of the possible categories, except for No melody, from which 10 series were selected. We included twice the number of series in the latter category in order to verify whether it contained tone series deemed melodies on grounds not accounted for by the perceptual mechanisms discovered thus far. The stimuli selected for the three fragment lengths are shown in Tables 1, 2, and 3.

Stimulus Presentation

The tone series of lengths 3, 4, and 5 were presented in three separate experimental sessions. All tone series were preceded by a cadence consisting of the chords C7-F to induce the key of F major. The cadence was always in a lower octave region (Figure 4). The interonset intervals between the two chords and between the tones in the series were 800 ms, as was the interval between the cadence and the tone series. In each trial, the pitch height of the stimulus (a cadence followed by a tone series) was determined randomly within a range of 6 semitones (except that two consecutive stimuli could never be presented at the same pitch). Moreover, the order of presentation was randomized independently for each subject. The tone series and the preceding chords were generated on a Rhodes 760 synthesizer, using a Piano sound (Acoustical Piano 5) and emitted via a Kawai KM-20 active speaker. Stimulus presentation and response collection were controlled by a special program running on an Atari 1040 STf computer. An example of a complete stimulus containing two chords and a series of tones is shown in Figure 4.

TABLE 1
Stimuli of Length 3 Used in Experiment 1

	Element			
Stimulus No.	1	2	3	Category*
1	G	C	E	Ch1
2	E	C	G	Ch1
3	C	G	E	Ch1
4	E	G	C	Ch1
5	C	E	G	Ch1
6	E	B♭	C	Ch2
7	B♭	E	C	Ch2
8	E	C	B♭	Ch2
9	C	B♭	E	Ch2
10	C	E	B♭	Ch2
11	C	F♯	G	A1
12	F♯	G	C	A1
13	E	F♯	G	A1
14	F♯	G	B♭	A1
15	B♭	F♯	G	A1
16	F♯	B♭	G	A2
17	F♯	C	G	A2
18	F♯	E	G	A2
19	C	F♯	B♭	N
20	E	C	F♯	N
21	C	B♭	F♯	N
22	F♯	C	E	N
23	E	F♯	C	N
24	C	E	F♯	N
25	B♭	F♯	C	N
26	B♭	C	F♯	N
27	F♯	C	B♭	N
28	F♯	B♭	C	N

* Last column indicates category to which the series belongs. Note that category A2 contains only 3 series, which is the maximum number available in that category. Ch1 = chord recognition C triad, Ch2 = chord recognition C7 chord, A1 = anchoring neighbor tone, A2 = anchoring with 1 intermediate tone, and N = no melody (does not form a melody).

Procedure

A button panel was used to control stimulus presentation and response indication. This panel contained 7 buttons: a start button to call up a stimulus, a repeat button to repeat a stimulus, and 5 buttons positioned in a row and labeled 1 to 5, serving as a 5-point scale (see Figure 5; the top button was not used in this experiment). Seven-hundred milliseconds after the start button was pressed, a stimulus (a 2-chord cadence followed by a tone series) was presented. The subjects received the following instruction, which was displayed on the computer screen:

> You will hear tone series of 3 (4, or 5, depending on the session) tones; each series will be preceded by two chords. For each series you must judge whether in your opinion it might occur in a normal (tonal) piece of music, or, in other words, whether it sounds good or bad in musical terms. Listen carefully to each series and then react spontaneously by pressing one of the buttons labeled 1 to 5 (1 = bad, 5 = good). Note that the series are not necessarily complete; in fact you will probably find most series incomplete.

TABLE 2
Stimuli of Length 4 Used in Experiment 1

	Element				
Stimulus No.	**1**	**2**	**3**	**4**	**Category***
1	E	G	B♭	C	Ch2
2	C	B♭	E	G	Ch2
3	C	G	E	B♭	Ch2
4	G	C	E	B♭	Ch2
5	E	C	B♭	G	Ch2
6	C	B♭	F♯	G	A1
7	B♭	C	F♯	G	A1
8	E	C	F♯	G	A1
9	E	B♭	F♯	G	A1
10	B♭	E	F♯	G	A1
11	B♭	F♯	E	G	A2
12	E	F♯	B♭	G	A2
13	C	F♯	B♭	G	A2
14	B♭	F♯	C	G	A2
15	E	F♯	C	G	A2
16	F♯	B♭	C	G	A3
17	F♯	E	B♭	G	A3
18	F♯	E	C	G	A3
19	F♯	C	E	G	A3
20	F♯	B♭	E	G	A3
21	E	G	C	F♯	N
22	C	F♯	B♭	E	N
23	C	E	F♯	B♭	N
24	B♭	F♯	C	E	N
25	C	B♭	E	F♯	N
26	C	E	B♭	F♯	N
27	F♯	C	B♭	E	N
28	E	C	F♯	B♭	N
29	C	E	G	F♯	N
30	F♯	B♭	C	E	N

* Last column indicates category to which the series belongs. Ch2 = chord recognition C7 chord, A1 = anchoring neighbor tone, A2 = anchoring with 1 intermediate tone, A3 = anchoring with 2 intermediate tones, and N = no melody (does not form a melody).

Pressing one of the scale buttons ended the trial. The following trial began by pressing the start button. A stimulus could be repeated (as long as the subject had not given a response) by pressing the repeat button. All sequences within a session were presented and responded to by using this self-paced routine, which allowed subjects to rest occasionally.

In order to familiarize subjects with the task, before beginning the experiment proper, the following practice scheme was employed. First subjects listened to 6 examples of so-called "good" series (series belonging to the categories Chord recognition 1, Chord recognition 2 or Anchoring 1, but not used in the experiment). After this, subjects listened to 6 examples of "bad" series (series belonging to the category No melody and also not used in the experiment). Two practice runs then followed that used 12 tone series selected from all categories used in the experiment but not belonging to the experimental stimuli. On the first run, subjects received feedback in the form "this series is usually judged good/bad"; on the second run, no feedback was given.

TABLE 3
Stimuli of Length 5 Used in Experiment 1

	Element					
Stimulus No.	1	2	3	4	5	Category*
1	B♭	E	C	F♯	G	A1
2	E	B♭	C	F♯	G	A1
3	E	C	B♭	F♯	G	A1
4	C	E	B♭	F♯	G	A1
5	C	B♭	E	F♯	G	A1
6	C	E	F♯	B♭	G	A2
7	E	F♯	C	G	B♭	A2
8	E	C	F♯	B♭	G	A2
9	C	B♭	F♯	E	G	A2
10	B♭	C	F♯	E	G	A2
11	F♯	C	E	G	B♭	A3
12	B♭	F♯	C	E	G	A3
13	E	F♯	B♭	C	G	A3
14	F♯	E	C	G	B♭	A3
15	C	F♯	B♭	E	G	A3
16	F♯	B♭	C	E	G	A4
17	F♯	C	B♭	E	G	A4
18	F♯	E	B♭	C	G	A4
19	F♯	B♭	E	C	G	A4
20	F♯	C	E	B♭	G	A4
21	C	G	F♯	E	B♭	N
22	C	E	G	F♯	B♭	N
23	G	C	E	B♭	F♯	N
24	B♭	G	F♯	E	C	N
25	G	F♯	E	C	B♭	N
26	E	G	C	F♯	B♭	N
27	C	G	E	B♭	F♯	N
28	G	B♭	C	F♯	E	N
29	G	F♯	C	E	B♭	N
30	B♭	G	C	F♯	E	N

* Last column indicates category to which the series belongs. A1 = anchoring neighbor tone, A2 = anchoring with 1 intermediate tone, A3 = anchoring with 2 intermediate tones, A4 = anchoring with 3 intermediate tones, and N = no melody (does not form a melody).

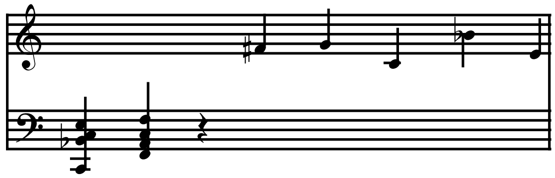


Fig. 4. An example of a stimulus consisting of a two-chord cadence followed by a tone series.

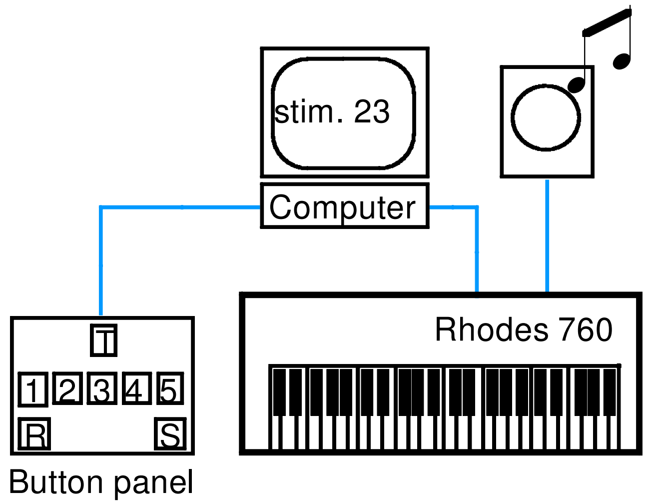


Fig. 5. Experimental setup. S = start button; R = repeat button; T = top button (Experiment 2b).

RESULTS

Consistency among respondents was determined by using an intraclass correlation coefficient (McGraw &Wong, 1996). For fragment sizes 3, 4, and 5, the intraclass correlation coefficients were .851, .842, and .617, respectively, all highly significant.

For each of the three stimulus sets of lengths 3, 4, and 5, we calculated the mean scale value for each stimulus averaged over subjects, as well as the mean scale value for the stimuli within each of the categories. Analyses of variance were then performed on each of the three sets of stimuli of lengths 3, 4, and 5.

The upper panel in Figure 6 shows the mean scale values obtained for the 5 categories of fragment length 3: Ch1, Ch2, A1, A2, and N. A mixed-model analysis of variance (ANOVA) showed a highly significant effect for the categories ($F(4, 76) = 24.09$, $p < .0001$). Post-hoc tests (Scheffé test; alpha = .05) indicated significant differences between the following pairs of categories: Ch1-A2, Ch2-A2, Ch1-N, Ch2-N, and A1-N.

The middle panel in Figure 6 shows the mean scale values obtained for the 5 categories of fragment length 4: Ch2, A1, A2, A3, and N. A mixed-model ANOVA showed a highly significant effect for the categories ($F(4, 76) = 39.77$, $p < .0001$). Post-hoc tests (Scheffé test; alpha = .05) indicated significant differences between the following pairs of categories: Ch2-A1, Ch2-A2, Ch2-A3, Ch2-N, A1-A2, A1-N, A2-N, and A3-N.

The bottom panel in Figure 6 shows the mean scale values obtained for the 5 categories of fragment length 5: A1, A2, A3, A4, and N. A mixed-

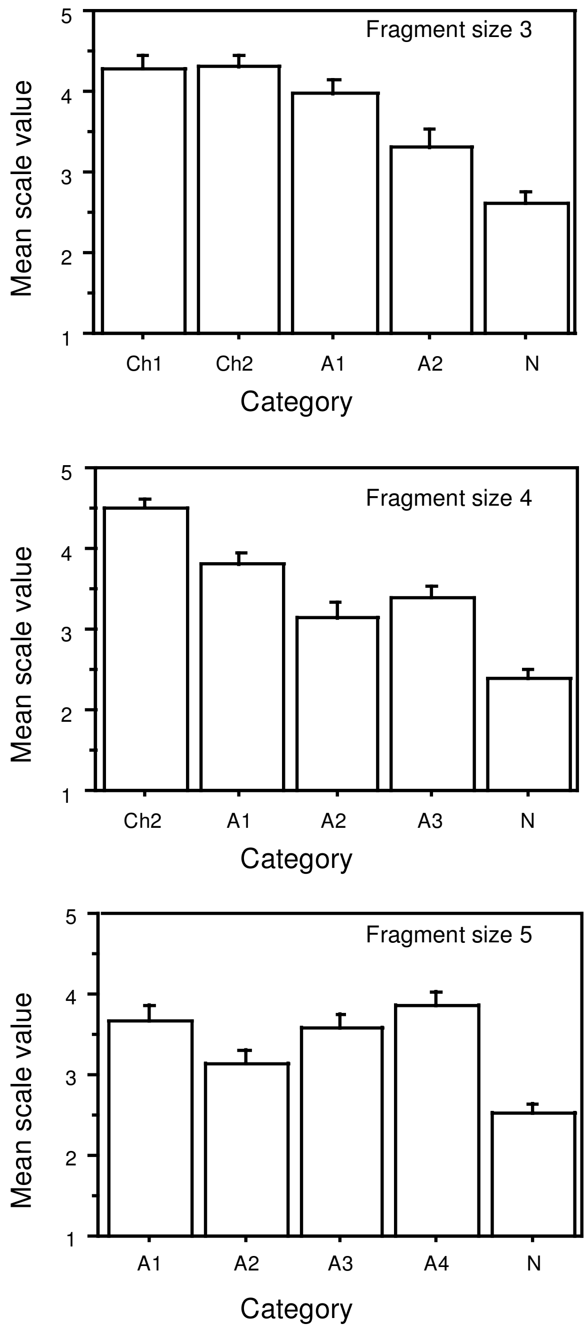


Fig. 6. Mean musical well-formedness ratings obtained for the stimuli in the different categories for the three stimulus sets of length 3, 4, and 5 of Experiment 1. Error Bars: ± 1 standard error. Ch1 = activation of a triad; Ch2 = activation of a seventh chord; A1 = immediate anchoring; A2, A3, A4 = anchoring with one, two, or three intermediate tones, respectively; N = does not form a percept.

model ANOVA showed a highly significant effect for the categories ($F(4, 76) = 17.437$, $p < .0001$). Post-hoc tests (Scheffé test; alpha = .05) indicated significant differences between the following pairs of categories: A1-N, A3-N, and A4-N. According to a Bonferroni/Dunn test, the difference between categories A2 and A4 is also significant.

DISCUSSION

The main result of this experiment is that percept formation of fragments of 3 to 5 tones, derived from the collection C_4 E_4 $F\sharp_4$ G_4 $B\flat_4$, can be explained by the hypothesized perceptual mechanisms: a tone series is judged a better (tonal) percept if one or more of the perceptual mechanisms is applicable than if the mechanisms are not applicable.

In particular, the three hypotheses stated above are supported by the following findings: (1) fragments containing only chord tones are generally rated highest; (2) fragments containing an F♯ followed by a G (either immediately or with intermediate tones) are generally rated lower; (3) fragments containing an F♯ never followed by a G are rated lowest.

Regarding the mechanism of anchoring, we tentatively hypothesized that melodic goodness would decrease with the number of intermediate tones between the anchored tone and its anchor, that is, that the judged goodness decreases from A1 through A4. Although judged goodness indeed decreases from A1 to A2, it *increases* from A2 to A3 and from A3 to A4. An explanation for this rather surprising finding may be sought in the location of the nonfitting F♯ in these stimuli (nonfitting in the scale of F to which all other tones belong), since it will either be in the first position (category A4; e.g., F♯ C B♭ E G) or in the second (category A3; e.g., B♭ F♯ C E G). This means that at the end of the presentation of the series, the three or four most recent tones can be captured in a musical percept. As it has been some time since the F♯ sounded, its memory trace may have decayed, as a result of which it does not contribute very much to the impression of the musical goodness of the entire series. This explanation implies that the perception (of the short sequences used in the present experiment) is rather local, taking into account only a few tones.

If the latter explanation is correct, it should apply equally to the stimuli in the N category. The stimuli in this category all contain an unresolved F♯, located at different positions in the series. Following the preceding reasoning, we ought to find a higher goodness rating for those series in which the F♯ appears earlier. The correlation between the location of the F♯ and the obtained rating was therefore computed, resulting in correlations of -.718 ($p < .05$), -.668 ($p < .05$), and -.568 ($p < .09$), df = 8, for the series of length 3, 4, and 5, respectively. Although these correlations are not very high, they do suggest that the activation of a note decays quite rapidly, supporting the idea that perception is relatively local.

Finally, we want to discuss two potential alternative explanations for the reported findings. First, we consider whether the principle of pitch proximity, mentioned in the Introduction, could explain the reported findings. Such an explanation would predict that sequences in which subsequent tones are close in pitch will form a musical percept. However, a proximity principle does not explain that series in which the F♯ precedes the G are judged good melodies, whereas series in which the G precedes the F♯ are judged bad.

Second, it might be argued that the sequences that contain only chord tones (Ch1, Ch2) receive higher ratings than the sequences also containing nonchord tones, simply because the former are diatonic melodies whereas the latter are not. This point of view does not need a chord recognition mechanism for its explanation and is therefore simpler. This explanation can indeed explain the different ratings for the Ch1, Ch2 sequences versus the others, but it cannot explain the different ratings in the sequences in which the position of the nonchord tone is varied. These differences can be explained only by taking into account both chord recognition and anchoring. Moreover the listeners' expectations for Ch1 and Ch2 fragments as observed in Experiment 2b appear to be resolutions of the chords, supporting the chord recognition hypothesis.

Experiment 2

Experiment 2 consisted of two sessions (a and b) in which listeners provided different responses for the same set of stimuli. Both sessions used a gating paradigm in which listeners were presented with stepwise lengthened fragments of tone sequences. This paradigm enables one to trace the development of a musical percept as the tone series unfolds. The principal purpose of this experiment is to test the hypothesis that the successive steps in the formation of musical percepts are determined by the applicability of the perceptual mechanisms chord recognition and anchoring.

Experiment 2a

In this session, listeners rated the melodic goodness of the incrementally presented fragments of a set of selected stimuli.

METHOD

Participants

Twenty musically experienced subjects, undergraduate and graduate students of the University of Nijmegen, participated in the experiment. All subjects were active music listeners, and the majority had played a musical instrument for several years (mean, 11.5

years). None had received formal music-theoretical training. The median age of the subjects was 24 years.

Stimulus Selection

Fourteen of the tone series used in the previous experiment were selected, namely those series of which all three fragments (length 3, 4, and 5) were presented. These 42 fragments are shown in Table 4. This selection allows us to compare the responses collected in the task environment of Experiment 1 with that of the present case.

Stimulus Presentation

As in the previous experiment, tone series were preceded by the chords C7 and F to induce the key of F major. The timing of the chords and the tones, the variation of the pitch height of the stimuli, and the randomization of stimulus order were identical to those in Experiment 1.

Procedure

The participants were seated in front of a computer and used a button panel to control stimulus presentation and to record the responses. The tone series were presented in a gating paradigm in which fragments of increasing size were presented. After pressing the start button, an initial fragment of length 3 of the series was presented, after which the subject judged the melodic goodness of the fragment on a five-point scale (by pressing the appropriate button on the panel). A stimulus could be repeated by pressing the repeat button. Pressing again on the start button, the subject would then hear the fragment of length 4. After having responded, the subject then listened to the last fragment of length 5. The paradigm is schematically shown in Figure 1. Working in this self-paced manner, the subject responded to all 14 series.

RESULTS

The intraclass correlation coefficient for consistency over subjects was .880, indicating high agreement among subjects. Next, mean scale values for the fragments in each of the 7 categories (Ch1, Ch2, A1, A2, A3, A4, and N) were calculated (Figure 7). A mixed-model ANOVA performed on these data showed a highly significant effect for the categories ($F(6, 114) = 16.37$, $p < .0001$). Post-hoc tests (Scheffé) indicated significant differences between the following pairs of categories: Ch1-A2, Ch1-N, Ch2-A2, and Ch2-N. According to a Bonferroni/Dunn post-hoc procedure, Ch1-A1, Ch1-A3, and A4-N also differ significantly.

Average scale values for the three fragments of each of the 14 stimuli were also calculated. These profiles (Figure 8) thus indicate the melodic goodness as judged for the successively increasing tone series. The category to which each fragment belongs is indicated by the labels Ch1, Ch2, A1, A2, A3, A4, and N.

DISCUSSION

We may conclude that the data support the hypothesis stated earlier that the development of (tonal) musical percepts is determined by the applicability of the mechanisms chord recognition and anchoring.

TABLE 4
Stimuli Used in Experiments 2a and 2b Clustered by Tone Series From Which They Are Derived

Stimulus No.	Element 1	2	3	4	5	Category in Expts. 1 and 2a*	Category in Expt. 2b*
1.3	G	C	E	—	—	Ch1	6
1.4	G	C	E	B♭	—	Ch2	6
1.5	G	C	E	B♭	F♯	N	1
2.3	C	G	E	—	—	Ch1	6
2.4	C	G	E	B♭	—	Ch2	6
2.5	C	G	E	B♭	F♯	N	1
3.3	C	E	G	—	—	Ch1	6
3.4	C	E	G	F♯	—	N	1
3.5	C	E	G	F♯	B♭	N	2
4.3	E	G	C	—	—	Ch1	6
4.4	E	G	C	F♯	—	N	1
4.5	E	G	C	F♯	B♭	N	2
5.3	C	B♭	E	—	—	Ch2	6
5.4	C	B♭	E	F♯	—	N	1
5.5	C	B♭	E	F♯	G	A1	4
6.3	C	E	B♭	—	—	Ch2	6
6.4	C	E	B♭	F♯	—	N	1
6.5	C	E	B♭	F♯	G	A1	4
7.3	E	F♯	C	—	—	N	2
7.4	E	F♯	C	G	—	A2	4
7.5	E	F♯	C	G	B♭	A2	4
8.3	F♯	C	E	—	—	N	3
8.4	F♯	C	E	G	—	A3	5
8.5	F♯	C	E	G	B♭	A3	5
9.3	C	E	F♯	—	—	N	1
9.4	C	E	F♯	B♭	—	N	2
9.5	C	E	F♯	B♭	G	A2	4
10.3	E	C	F♯	—	—	N	1
10.4	E	C	F♯	B♭	—	N	2
10.5	E	C	F♯	B♭	G	A2	4
11.3	B♭	F♯	C	—	—	N	2
11.4	B♭	F♯	C	E	—	N	3
11.5	B♭	F♯	C	E	G	A3	5
12.3	C	F♯	B♭	—	—	N	2
12.4	C	F♯	B♭	E	—	N	3
12.5	C	F♯	B♭	E	G	A3	5
13.3	F♯	B♭	C	—	—	N	3
13.4	F♯	B♭	C	E	—	N	3
13.5	F♯	B♭	C	E	G	A4	5
14.3	F♯	C	B♭	—	—	N	3
14.4	F♯	C	B♭	E	—	N	3
14.5	F♯	C	B♭	E	G	A4	5

* These categories are defined in the text.

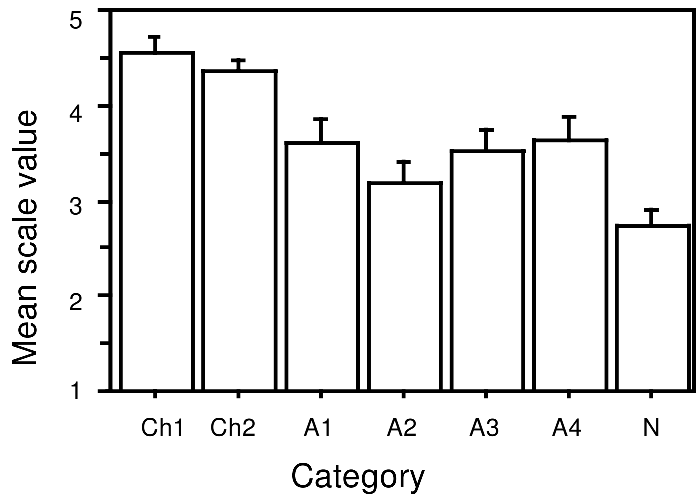
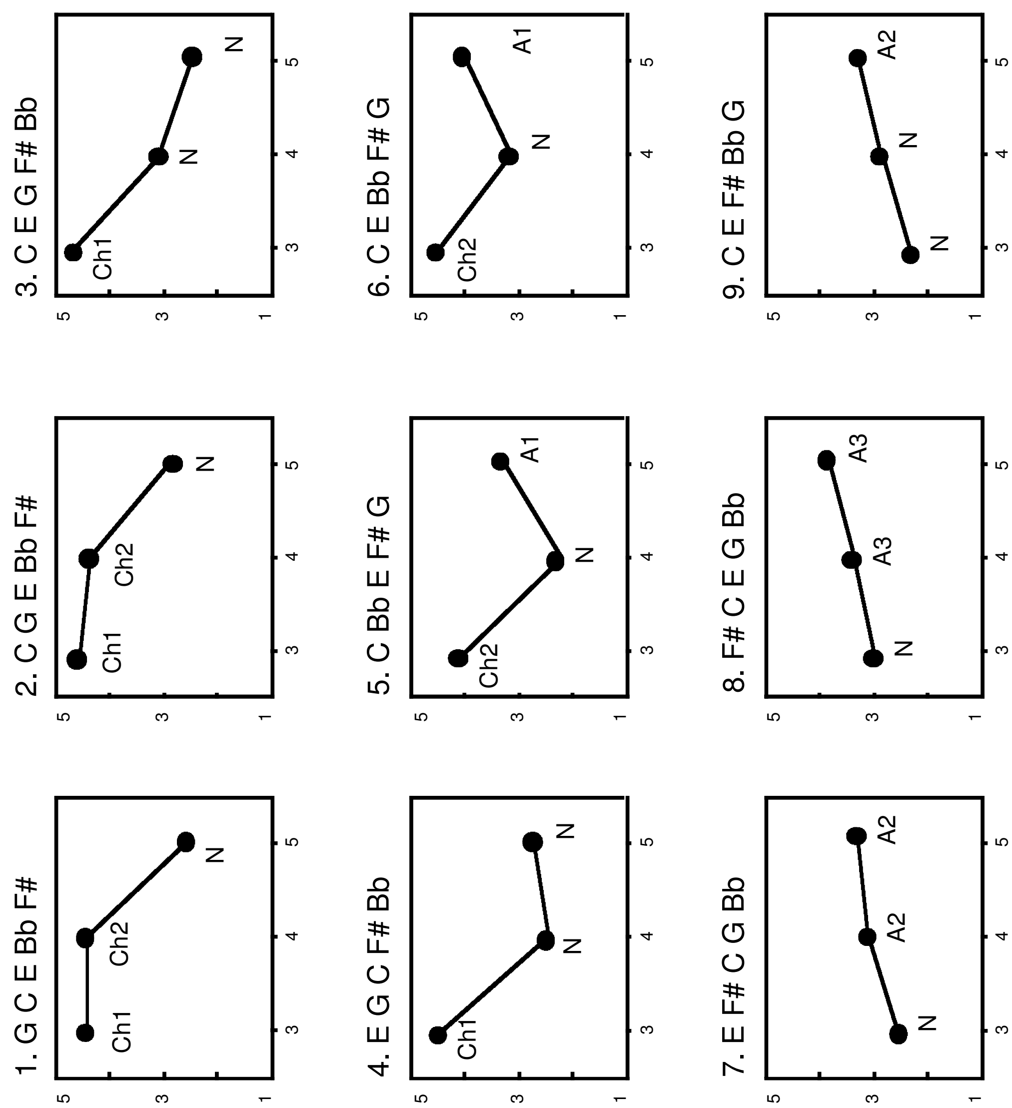


Fig. 7. Mean musical well-formedness ratings for the fragments in the 7 categories obtained in Experiment 2a. Ch1 = activation of a triad; Ch2 = activation of a seventh chord; A1 = immediate anchoring; A2, A3, A4 = anchoring with one, two, or three intermediate tones, respectively; N = does not form a percept.

First, the histogram in Figure 7, displaying the accumulated goodness judgments across fragments of different lengths presented in the gating paradigm, shows that fragments of type Ch1 and Ch2 (inducing a triad and a dominant-seventh chord, respectively) receive the highest rating for melodic goodness, the ratings for fragments of type A1 and A2 (direct and delayed anchoring) are lower, those for the type A3 and A4 series receive a somewhat higher rating, and the fragments of type N receive the lowest rating, just as in the Experiment 1. This result indicates that the perceptual mechanisms of chord recognition and anchoring also explain percept formation across different series lengths.

Second, the profiles in Figure 8, displaying the melodic goodness ratings for each of the three fragments of the 14 tone series, indicate that judged goodness strongly fluctuates from one fragment to the next. The rating of each successive fragment depends on the applicability of perceptual mechanisms that transform the fragment into a musical percept. Ratings of fragments of the same type do sometimes differ, for example, the ratings of the N fragments in stimuli 10 to 14, which we cannot as yet explain. Overall, the findings support the hypothesis that the time course of percept formation is guided by the mechanisms of chord recognition and anchoring.


1. G C E Bb F#
2. C G E Bb F#
3. C E G F# Bb
4. E G C F# Bb
5. C Bb E F# G
6. C E Bb F# G
7. E F# C G Bb
8. F# C E G Bb
9. C E F# Bb G
Rating
Ch1
Ch2
N
A1
A2
A3
1
3
5
3
4
5

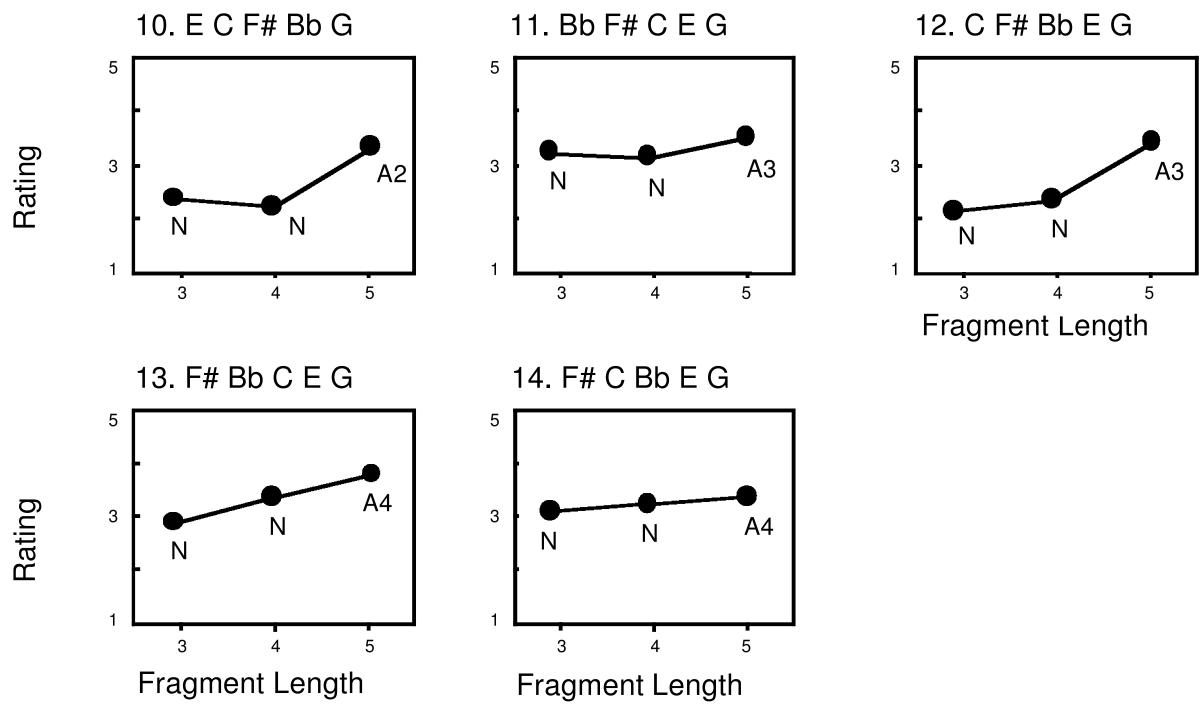


Fig. 8. Mean musical well-formedness ratings for initial fragments of increasing size (length 3, 4, and 5) for the 14 tone series of Experiment 2a.

Experiment 2b

Experiments 1 and 2a indicated that the mechanisms of chord recognition and anchoring could explain both the perception of short tone sequences (Experiment 1) and the development of percept formation of stepwise increasing fragments (Experiment 2a). In the present session, listeners generated continuations for the tone series of Experiment 2a, which may provide an independent and more direct source of evidence concerning the hypothesized perceptual mechanisms.

METHOD

Participants

Eighteen subjects took part in the experiment, all of whom had also participated in the previous experiment. These 18 subjects all played a keyboard instrument (mainly piano) and were able to find a series of tones on the synthesizer.

Stimulus Selection and Presentation

The same stimuli as in Experiment 2a were used. The method of stimulus presentation, the gating paradigm, the preceding cadence, the timing, and the randomization of stimuli were also the same as in Experiment 2a. The task for the participants, however, was different.

Procedure

After having been presented with a stimulus, each subject was asked, if feasible, to sing a few tones that complete the series (maximum 5). Next, the subject was asked to find these completing tones on the synthesizer. During this search process, the stimulus could be repeated by pressing the repeat button. When the tones had been found, the subject pressed the top button on the button panel and played the completing tone(s) once more, after which the subject pressed the highest key on the synthesizer, which indicated to the computer that the intake could be stopped. If the subject was unable to continue the series, (s)he proceeded to the next stimulus by pressing the start button (see Figure 5). This rather complicated procedure was necessary in order to guarantee an error-free registration of the completion series. Before the experiment, the subjects trained with a few series not used in the experiment.

RESULTS

The completions generated by the subjects were collected for each of the three initial fragments of the 14 stimuli presented. For 95 of the 756 fragments (18 subjects × 14 stimuli × 3 fragments) presented, subjects were unable to complete the series. Sixty-five of those cases pertained to a tone series in the N category (no melody), suggesting that the main reason for not being able to complete the series was its absence of a musical percept. The number of tones used in the continuations varied between 1 and 5, with a mean of 2.02 tones. No systematic relation seemed to exist between type of series and the number of tones used to complete it.

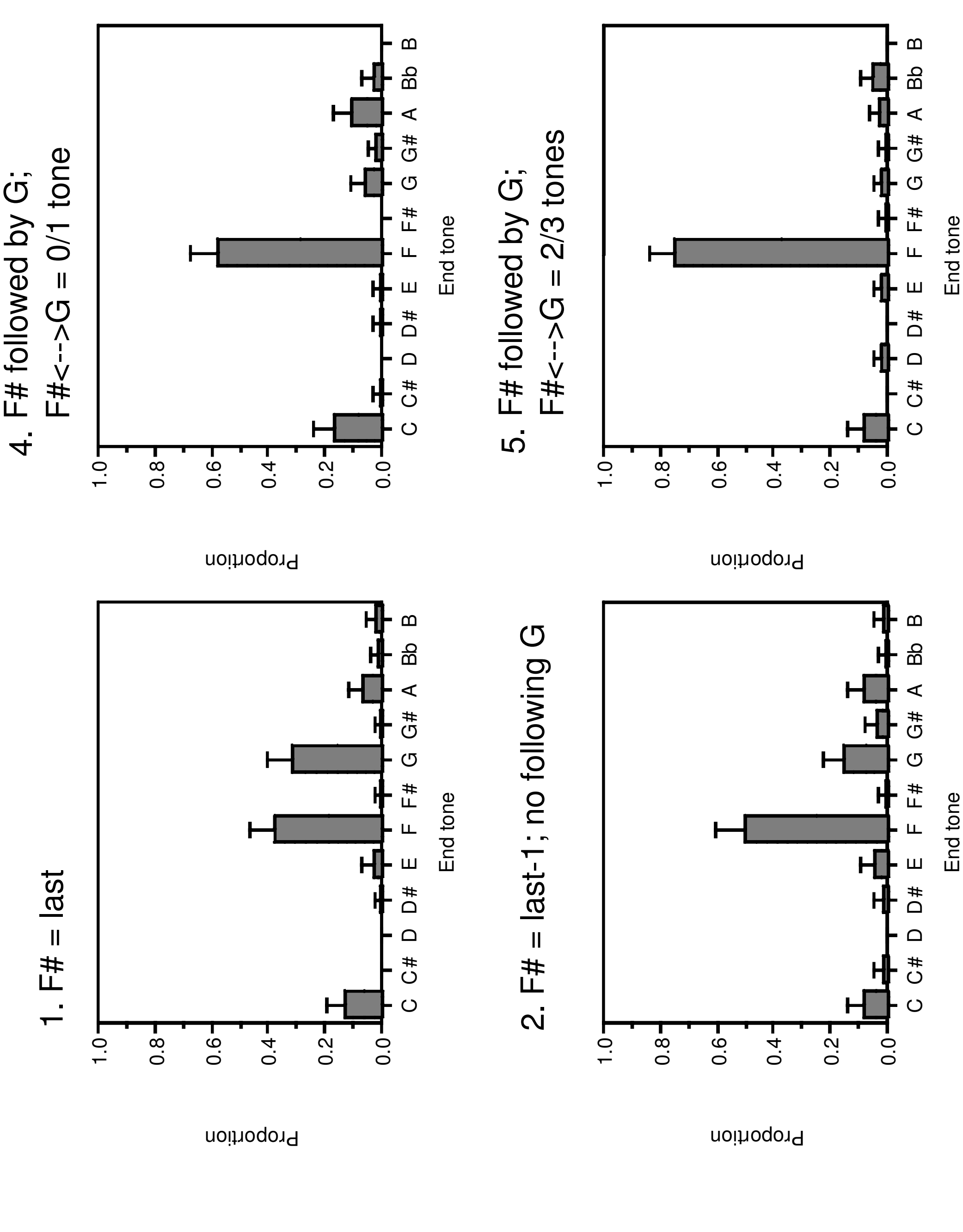
Frequency distributions of the continuations for each of the 42 fragments were made, with the results being pooled over subjects. Two different distributions were calculated: one for the end tones of the completions and one for all other tones (called the transient tones). The 42 fragments were then divided into 6 categories defined by specific characteristics of the tone series, as follows: (1) Last tone is an F♯ (comprising 8 fragments); (2) Penultimate tone is an F♯ that is not followed by a G (7 fragments); (3) Either the second from last tone or the third from last tone is an F♯ that is not followed by a G (7 fragments); (4) The series contains an F♯ that is either immediately followed by a G or with one intermediate tone (6 fragments; this category includes the categories A1 and A2 of Experiment 1); (5) The series contains an F♯ that is followed after two or three tones by a G (6 fragments; this category includes the categories A3 and A4 of Experiment 1); (6) The series does not include an F♯ (8 fragments; this category includes the categories Ch1 and Ch2 of Experiment 1). The last column of Table 4 indicates to which category each fragment belongs. Frequency distributions for the end tones and the transient tones are shown in Figures 9 and 10, respectively. As an indication of the reliability of the subjects' expectations, we calculated the 95% confidence intervals, which are added to the histograms as error bars.

DISCUSSION

The main purpose of Experiment 2b was to collect independent evidence in support of the hypothesized perceptual mechanisms examined in Experiments 1 and 2a. When a subject senses that a fragment can be completed, it means first that a musical percept has been formed and second that the percept contains one or more "unresolved tendencies" that will shape its continuation. Thus by examining the continuations we can discover the unresolved tendencies, which in turn may reveal the created percepts. This may be conceived as follows.

The mechanism of chord recognition posits the activation of the chord C or C7 (categories Ch1 and Ch2, respectively), which in the context induced will both be perceived as chords on the dominant, resulting in an expectation for an element of the F triad, preferably the tonic F itself. Because F is the central tone of the induced key, it will most likely be the *final* tone of the continuation.

The mechanism of anchoring posits that the tone F♯, within the context induced, will resolve to G (direct anchoring (A1), or delayed anchoring (A2)). The mechanism of anchoring in fact assumes that the tone F♯ creates an expectation for the tone G, irrespective of its actual subsequent occur-


1. F# = last
2. F# = last-1; no following G
4. F# followed by G;
F#<-->G = 0/1 tone
5. F# followed by G;
F#<-->G = 2/3 tones
Proportion
End tone
C C# D D# E F F# G G# A Bb B
0.0 0.2 0.4 0.6 0.8 1.0

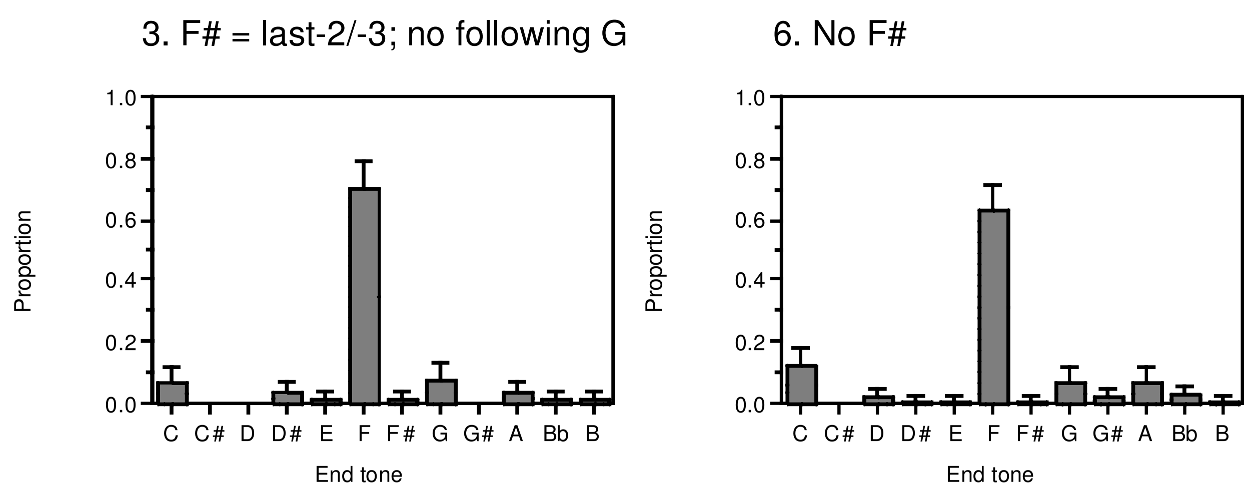
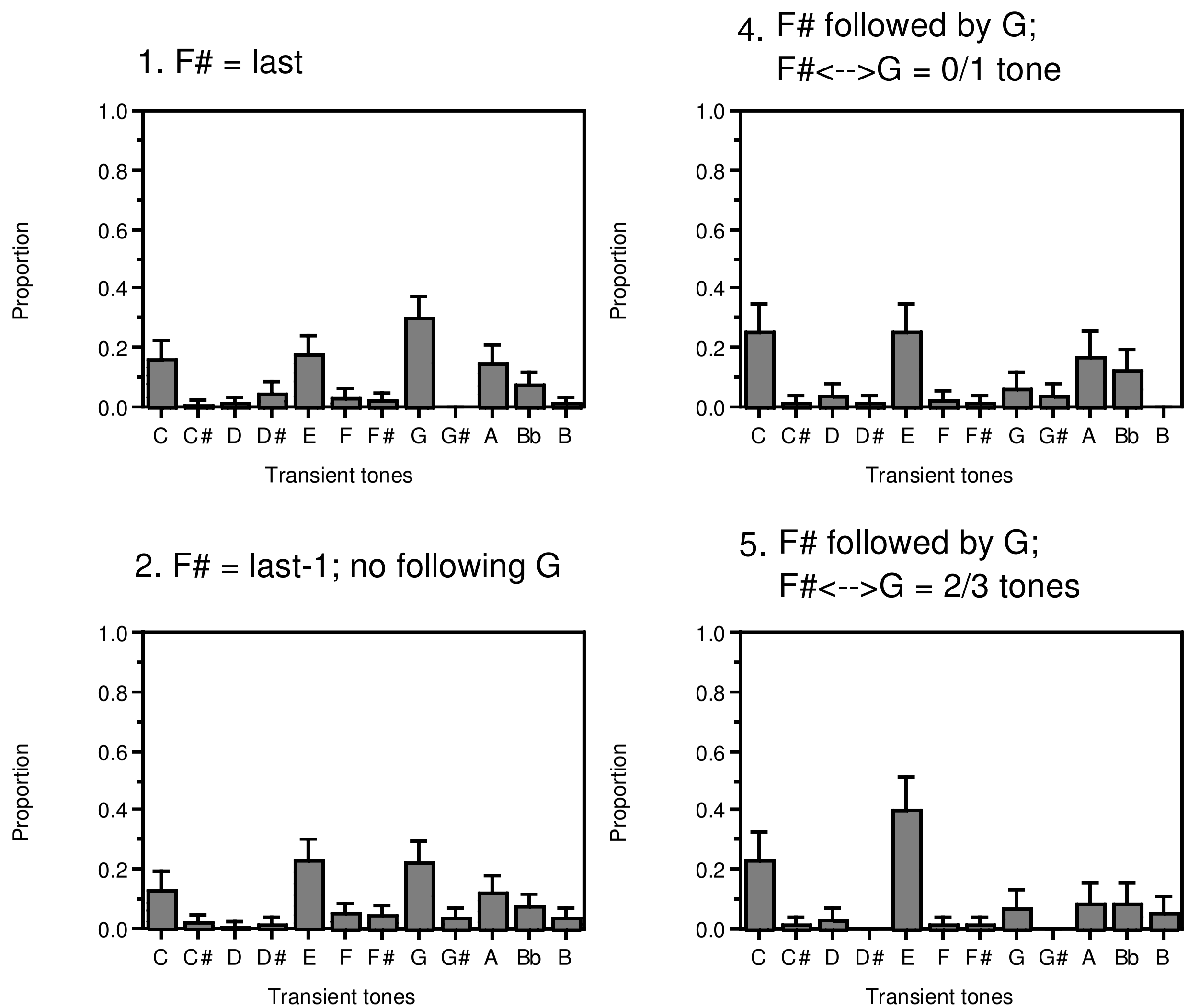


Fig. 9. Frequency distribution of the end tones in the complements generated in response to the stimuli in the 6 categories of Experiment 2b. Category definition is specified on top of each histogram. Error bars indicate 95% confidence intervals.



1. F# = last
2. F# = last-1; no following G
4. F# followed by G;
F#<-->G = 0/1 tone
5. F# followed by G;
F#<-->G = 2/3 tones
Proportion
1.0
0.8
0.6
0.4
0.2
0.0
C C# D D# E F F# G G# A Bb B
Transient tones

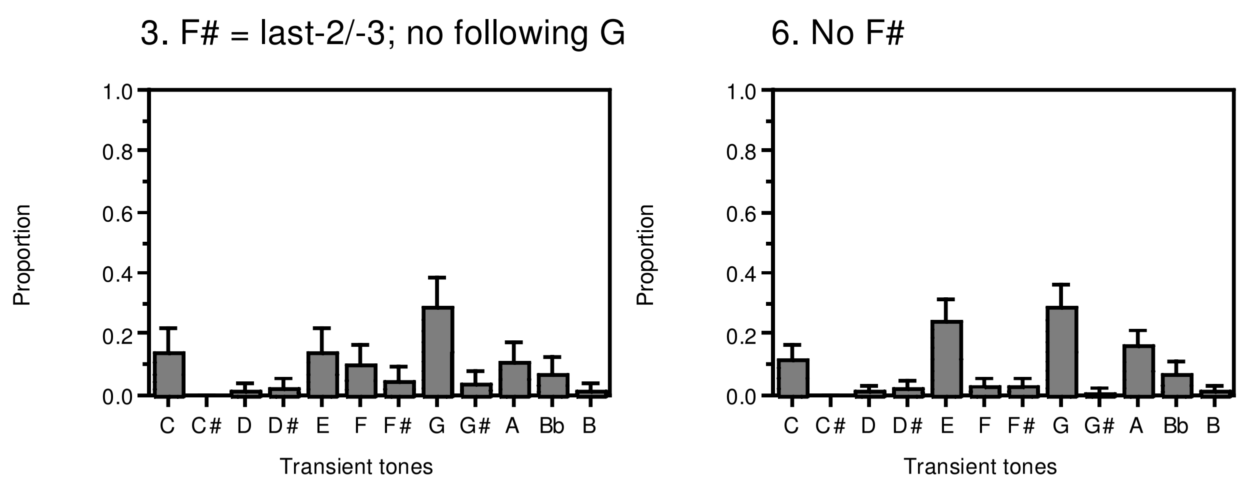


Fig. 10. Frequency distribution of the transient tones in the complements generated in response to the stimuli in the 6 categories of Experiment 2b. Category definition is specified on top of each histogram. Error bars indicate 95% confidence intervals.

rence. If the tone G does occur, anchoring will take place. But if the F♯ is not followed by a G, an expectation for its occurrence will still arise, which will be expressed by the appearance of a G in the continuation, probably not as a final tone but as a *transient* tone.

When the frequency distributions of the final and the transient tones are compared, it may be observed that the F clearly dominates among the end tones (Figure 9), whereas it hardly occurs among the transient tones (Figure 10). This finding supports the supposition that the tones C, E, G, and B♭ activate the dominant chord C7, a chord with a strong tendency to resolve to F. The fact that the F hardly occurs among the transient tones supports the idea that the F functions as a goal tone for the completion.

Considering the distribution of end tones, we observe a relatively high frequency of Gs among the continuations for the fragments in Category 1. This can be understood as resulting from the recent occurrence of the tone F♯, which is associated with an expectation for a succeeding G. This finding supports the basis of the hypothesized anchoring mechanism: the occurrence of an F♯ creates an expectation for a G. If the G actually occurs, there is anchoring; if it does not, the G is generated in the completion.

Looking next at the distribution of the end tones in the Categories 1, 2, and 3, we see that the more tones that have sounded before the occurrence of the tone F♯, the smaller the number of Gs. A similar pattern is found in the distribution of the transient tones. This may be understood as the result of a decreasing activation of the tone F♯ (caused either by decay or by interference with the tones succeeding the F♯) and a concomitant decrease in the expectation of a G.

Also the decrease in the number of Gs among the end tones in moving from Category 1 to 3 is accompanied by an increase in the number of Fs. This increase can be understood as resulting from an interference between the two existing unresolved tendencies: a tendency associated with the tones of the dominant chord to generate an F, and a tendency associated with the F♯ to generate a G. The more recent the F♯, the greater the tendency to produce a G as the *final* tone, thereby "forgetting" the still existing tendency to produce a final F. That both tendencies do exist can be inferred by comparing the frequency distributions for the transient tones with the final tones of Category 1: there is a very high frequency of Gs in the transient tones and some 40% of Fs in the final tones, indicating that both tendencies are resolved in the listeners' responses.

Last, coming to the transient tones in the completions, we see that the tones C, E, G, and A occur most frequently. Except for the occurrence of the G, which may result from the resolution of a preceding F♯ as just argued, the occurrence of the other tones is probably determined not by unresolved tendencies but by melodic and rhythmic factors. That is, the generation of a completion for a presented tone series may be seen as a compositional process in which the resolving tones must be placed in a musical context that fits in with that of the presented series. Thus the subject will tend to take into account both melodic (e.g., preferably proceed by step) and rhythmic (e.g., the completion must conform with the grouping of the tone series) principles. The transient tones C, E, and A thus contribute little to a better understanding of the percepts formed, for which reason we have not analyzed them further.

General Discussion

In three experimental sessions, we examined the on-line processing of brief tone sequences in the Western tonal system. In all three cases, (initial) fragments derived from different orderings of the sequence C_4 E_4 $F\sharp_4$ G_4 $B\flat_4$ were used. In Experiment 1, those fragments to which the perceptual mechanisms chord recognition and anchoring apply yielded significantly better percepts than did those fragments to which these mechanisms do not apply. It should be noted that chord recognition and anchoring are distinct perceptual mechanisms. Chord recognition is the result of a successful memory search in which a tone series is recognized as a pattern stored in long-term memory. This process is comparable to that by which a word is recognized in a stream of speech sounds. Anchoring, however, is not an instance of recognizing a pattern stored in memory, but results from the detection of a structural relation between elements in the input, thereby relying on knowledge about how music is created. In the case of anchoring, a tone is identified as a relatively unstable element (the F♯ in the sequences presented) that may be linked to a more stable tone close in pitch and time (the G), by which the unstable element is integrated in the percept.

In Experiment 2a, we examined the development of a musical percept when subjects listen to fragments of increasing length and then indicate their melodic goodness. The percepts formed for fragments of differing length can also be understood as resulting from the application of the perceptual mechanisms chord recognition and anchoring. The results show that during the incremental buildup of a musical percept that occurs while listening to an unfolding tone series, well-formed and less well-formed percepts may alternate, depending on whether the current fragment can or cannot be described by means of the perceptual mechanisms. Thus this experiment has provided some insights into the incremental aspects of music perception by revealing the listener's continual attempt to produce a complete description of the fragment heard thus far, an attempt that may or may not be successful.

If we combine the results from this experiment with the finding, suggested by the results of Experiment 1, that the listener's perceptual window contains only a few tones, or in other words, that the impression of past tones decays relatively rapidly, then there emerges a picture of music processing that is oriented rather locally and aimed at making a coherent description of the most recent few tones within a frame of interpretation evoked by the larger context.

In search for further evidence to support the hypothesized perceptual mechanisms, in Experiment 2b we examined the continuations, or more precisely the completions that listeners generate for the same fragments used in Experiment 2a. We found that if a fragment is describable as either a C triad or as a C7 chord, both functioning as a dominant chord in the induced key context, then the last tone of the completion was in almost all cases F, thus supporting the hypothesis that the series was interpreted as a dominant chord. Similarly, if a fragment contained an F♯ not resolved by a subsequent G, then a great number of Gs appeared in the continuation, confirming the hypothesis that, in the context induced, the F♯ is experienced as an unstable tone that tends to be resolved by G.

In sum, we may conclude that this study has provided converging evidence that the percept formation of the tone series examined is determined by the operation of the two perceptual mechanisms, chord recognition and anchoring. Apart from these preliminary results, we believe that the paradigm in which subjects listen to tone series of increasing length and indicate the melodic goodness of the series, or complete the series by a few tones, will serve as an effective tool in revealing the perceptual mechanisms underlying the processing of tone series.[2,3]

References

Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16*, 485–518.

Bharucha, J. J. (1996). Melodic anchoring. *Music Perception, 13,* 383–401.

Bigand, E. (1990). Abstraction of two forms of underlying structure in a tonal melody. *Psychology of Music, 18,* 45–60.

Cooke, D. (1959). *The Language of Music*. Oxford: Oxford University Press.

Cuddy, L. L., Cohen, A. J., & Mewhort, D. J. K. (1981). Perception of structure in short melodic sequences. *Journal of Experimental Psychology, Human Perception and Performance, 7,* 869–883.

Deliège, I. (1992). Recognition of the Wagnerian Leitmotiv. *Jahrbuch der Deutschen Gesellschaft für Musikpsychologie, 9,* 25–54.

Deliège, I. (1996). Cue abstraction as a component of categorisation processes in music listening. *Psychology of Music, 24,* 131–156.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception and Psychophysics, 28,* 381–389.

2. Part of this research was reported at the Third Triennial ESCOM Conference, Uppsala, Sweden, 7-12 June 1997.

3. We thank René van Egmond and the anonymous reviewers for their insightful comments, which greatly helped to improve the manuscript.

Deutsch, D. (1982). Grouping mechanisms in music. In D. Deutsch (Ed.), *The psychology of music*. San Diego, CA: Academic Press.
Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences. *Psychological Review, 88*, 503–522.
Halpern, A. (1984). Perception of structure in novel music. *Memory & Cognition, 12,* 163–170.
Howell, P., West, R., & Cross, I. (1991). *Representing musical structure*. London: Academic Press.
Krumhansl, C. L., (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.
Laden, B. (1994). Melodic anchoring and tone duration. *Music Perception*, *12*, 199–212.
Lerdahl, F. (1988). Tonal pitch space. *Music Perception, 5,* 315–350.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30–46.
Povel, D. J. (1996). Exploring the elementary harmonic forces in the tonal system. *Psychological Research, 58,* 274–283.
Povel, D. J. (1997). Processes in on-line music perception. *Proceedings of The Third Triennial ESCOM Conference,* 251–256. Uppsala, Sweden: ESCOM.
Povel, D. J., & Van Egmond, R. (1993). The function of accompanying chords in the recognition of melodic fragments. *Music Perception, 11,* 101–115.
Serafine, M. L., Glassman, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. *Music Perception, 6,* 397–430.
Sloboda, J. A., & Parker, D. (1985). Immediate recall of melodies. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 143–168). London: Academic Press.
Tan, N., Aiello, R., & Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Memory and Cognition, 9,* 533–539.
Zuckerkandl, V. (1956). *Sound and symbol*. Princeton, NJ: Princeton University Press.