DETECTING AND RESOLVING METRICAL AMBIGUITY IN A ROCK SONG UPON MULTIPLE REHEARINGS

Peter Vazan & Michael F. Schober

New School for Social Research

ABSTRACT

This study examined listeners' detection and resolution of metrical ambiguity in the song 'Murder by Numbers' by The Police over the course of five hearings. In the extended opening of the song, a syncopated pattern in the drums ostensibly projects a triple meter that clashes with a duple meter played by the instruments that enter later. A tapping task was used to examine changes in listeners' perception of the ambiguous segment on multiple rehearings.

All participants first heard the whole song (Trial 1). In the remaining four trials participants were divided into two experimental groups receiving different instructions. Participants in the *natural* condition continued to tap the beat they spontaneously perceived in the subsequent presentations of the 80-second-long target segment. Participants in the *coherent* condition were asked to find one coherent way of tapping that would fit the whole segment.

The majority of listeners processed the musical signal primarily in a bottom-up fashion, interpreting the song as containing a change of meter even during the later trials. Only in a few cases did participants recognize the garden path nature of the segment and inhibit their initial hearing, reinterpreting the meter by applying some type of top-down listening strategy.

This study provides initial evidence on how interpretation of ambiguous meter changes upon rehearing. Although the findings suggest that the coherence (global stability) of metrical structure is not of primary importance for many average listeners, nonetheless there are trends toward greater coherence over repeated listening, particularly for some listeners.

1. INTRODUCTION

When listeners encounter a segment of music that allows alternate metrical interpretations, what leads them to choose one interpretation over another? If subsequently it becomes evident that their initial interpretation didn't fit with what came later, what do listeners do? Do they reinterpret what came before, or decide that the meter has changed? And how do these interpretations change upon subsequent hearings?

Some theories of ambiguity resolution in music [5, 7] propose that listeners compute several analyses of a signal simultaneously and at the moment of disambiguation choose the most fitting one, which is retrospectively projected backward to the beginning of the ambiguous passage. Such theories closely follow a view of reinterpretation in language that assumes that reinterpretation is very quick, automatic, and guided by a sense of coherence of the grammatical structure as well as coherence of meaning.

These views aren't the only possibilities. It is not necessarily the case that listeners assume coherence of musical structure in the same way they do in language understanding. And different kinds and degrees of ambiguity may lead listeners to different modes of reinterpretation. Just as in linguistic interpretation, if one metrical interpretation in music sounds much more natural to a listener, that is, more stable and robust than other alternatives, the music may not sound ambiguous at all.

We propose that people hearing a metrically ambiguous segment of music will follow alternate routes and hear the music from a different perspective that is *locally* less stable only when they want to achieve a *globally* coherent hearing of the whole piece or its larger segments. Listeners who aren't worried about whether the piece is metrically coherent will be likely to conclude that a metrically ambiguous piece changes meter. That is, they will process the music primarily bottom-up, seeking local stability by adjusting or changing reference frames (meter), without considering what a globally stable hearing would require. As a result, syncopated events can become desyncopated and off-beats can be heard as on-beats. Alternatively, a listener who does care about global metrical coherence will use a top-down process, taking into account broader context and being willing to sacrifice local stability.

Which approach are listeners more likely to take? Prior evidence from studies on the perception of (large-scale) tonal closure [2, 4, 9—Experiment 2] suggests that people's hearing consists of a series of more or less fragmented psychological effects rather than of a perception of a wholly integrated compositional structure[1]. But less work that directly addresses our questions has been done in the temporal domain. Although several studies have examined the differential processing of stable (on-beat) and unstable (off-beat) rhythmic events with respect to an underlying metrical framework [e.g., 3, 10, 11, 12, 18], most of them have not used real music as experimental stimuli, nor have they been concerned with the issue of ambiguity resolution.

The problem of ambiguity has been tackled more directly by Jackendoff (1991), who argues for a parallel multiple-analysis model which, as opposed to serial models, allows for a comparison between alternative analyses computed concurrently. Jackendoff contends that because of the parallel processing listeners can suddenly switch between analyses, instantly reinterpreting ambiguous stimuli. He supports this notion with evidence from psycholinguistic studies: as Swinney [14, 15], Tanenhaus, Leiman, and Seidenberg [16], and many others have demonstrated, for a brief time after a word is heard in a spoken sentence all of its meanings are active regardless of preceding semantic and syntactic context. Jackendoff argues that given the similarity of the tasks, the evidence for a parallel multiple-analysis model of language processing bolsters the plausibility of the parallel model for musical processing. However, since these ideas have not been experimentally tested in a real music listening situation, they remain hypothetical. The current study tests these ideas empirically.

2. EXPERIMENT

Stimulus description. The experimental stimulus was the rock song 'Murder by Numbers' by The Police, which contains extensive metrical ambiguity. The song has qualities similar to "garden path" sentences, which initially mislead listeners (or readers) with syntactic or lexical ambiguities and only later evidence forces them to reanalyze. Likewise, in the first 40 seconds of 'Murder by Numbers' a triple meter is strongly projected by a drum pattern, only to be suppressed when other instruments enter, projecting a duple meter that does not change for the rest of the song. The beginning of the song, however, can also be accommodated within a duple framework—that is, listeners could hear the drums as playing a highly syncopated rhythm, so as to arrive at a coherent and unified hearing of the whole song from its very first moment. It is the duple meter in which the song was composed and which can be found notated in a published score [13].

Figure 1 shows the introductory drum part notated from two metrical perspectives. In the beginning of the song, starting with the drums alone, the salience—that is, the acoustic strength—of the bass-drum and the rim shots on a snare drum projects a 3/4 (triple) meter with the bass-drum in the downbeat position. As a consequence, cymbals and hi-hat strokes are heard predominantly in off-beat positions (Figure 1*a*). However, later development in the piece with the overriding influence of chord changes, that is, the harmonic rhythm (along with the melody and the final disambiguation in the chorus), disconfirms the triple hearing and reveals a 12/8 meter—that is, a 4/4 duple meter[2] with triple subdivisions. In this meter the bass-drum and the rim shots function as the components of a rhythmic figure in which the bass drum marks off the second (and the fourth) beat of the tactus—the most salient level of beats people tend to clap along with—while the splash cymbal and hi-hat strokes take on strong beat positions (Figure 1*b*). In sum, what eventually turns out to be a duple meter that begins on a downbeat starts out sounding like a triple meter that begins with an anacrusis.

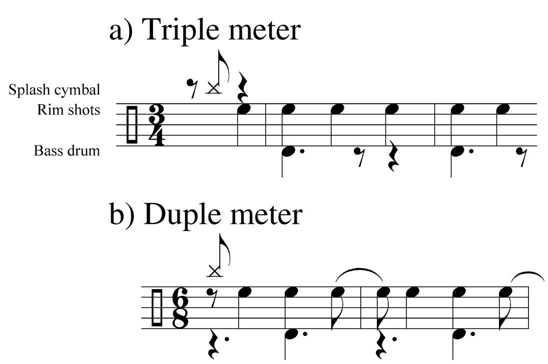


Fig. **1**. Two notated versions of the introductory drum-set part from 'Murder by Numbers' by The Police

Figure 2 graphically lays out the schematic representation of the ambiguity of this musical situation. In the triple meter interpretation, the rhythmic or durational value of the tactus beats is a quarter note, as projected by the most salient cues in the music. In the duple meter interpretation it is a dotted quarter note, a durational value projected by less salient cues. This means that the tempo of the duple interpretation is 50 percent slower than that of the triple interpretation; the difference amounts approximately to a 600 vs. a 400 *ms* interbeat interval (IBI) per tactus beat (which corresponds to 100 vs. 150 BPM). Moreover, the difference between the position of the duple meter downbeat and that of the following triple meter downbeat amounts to a time span of three eighth-notes (which in terms of absolute time corresponds to approx. 600 *ms*).

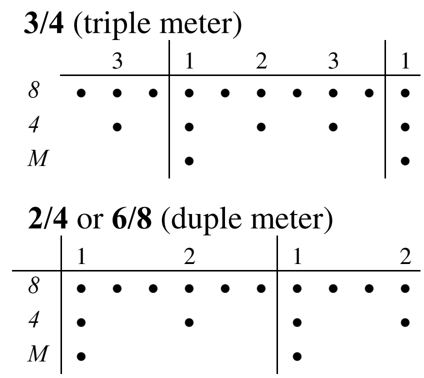


Fig. **2**. A pair of metrical grids revealing the structural features of the metrical ambiguity in the song

In sum, the prediction based on the music-theoretical analysis is that the triple interpretation is perceptually more stable at the very beginning of the song. But this interpretation does not lead to a globally stable and coherent hearing of the whole song. Global stability may be achieved by choosing a locally and initially less stable duple interpretation. Such an option is likely to be discovered only after one's hearing has been initially misled down the garden path by a hearing in triple meter. Thus, it is of particular interest to compare the initial interpretation and subsequent hearing of the very same stimulus. Figure 3 shows three progressively stronger instances of duple meter cues (which are the vocal melody, entrance of harmony and the timing of chord changes, and finally the unambiguous duple character of the chorus) that appear at the beginnings of and throughout Sections 2, 3, and 4. These are the main factors that may lead to disambiguation of the meter during the initial hearing of the song and trigger *retrospective rehearing* in the sense proposed by Jackendoff.

One condition of the experiment asked whether people spontaneously discover the presence of the metrical ambiguity, and if so, whether on subsequent presentations they attempt to arrive at a coherent hearing through reinterpretation of the drum introduction. People who were assigned to this *natural* condition were simply asked to tap the meter (the regular train of beats) they spontaneously felt or heard in each of the five trials. The main purpose of this condition was to see how people naturally hear the song without experimental instructions to listen in any particular way. Another objective was to examine whether the detection of ambiguity and the time course of the reanalysis in music is similar to that

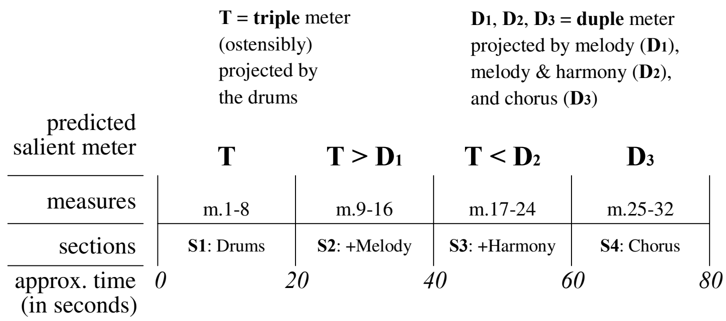


Fig. 3. Structure of the opening segment from ‘Murder by Numbers’ with an indication of relative salience of both meters in each of the four sections

which has been observed in language processing. In a second condition, termed the *coherent* condition, participants were informed in advance that the song contains metrical ambiguity and they were instructed to try to arrive at one coherent interpretation of it. However, they were not informed whether coherent tapping meant tapping in duple or triple meter. Accordingly, our second question was whether explicit instruction to arrive at a coherent interpretation would affect rehearings. The idea was that, even if people’s natural interpretation is that the song changes meter (a bottom-up strategy), their listening strategies may differ substantially if the ambiguity is called to their attention.

In both conditions, the following tapping scenarios were possible: (*a*) a participant starts to tap in triple meter and continues tapping in this manner (disregarding the disambiguation), even on subsequent hearings; (*b*) a participant changes his or her tapping, switching from triple to duple meter, thus interpreting the song as changing the meter at a certain point—the precise point might or might not shift on subsequent hearings; (*c*) a participant initially taps in triple meter, but in subsequent hearings attempts to impose duple meter on the introductory ambiguous segment of the song; (*d*) a participant taps in duple meter from the very first moment and keeps tapping that way.

If bottom-up processing prevails, listeners will either not detect the ambiguity at all or if so they will not resolve it. If top-down processing is stronger, listeners may establish structural coherence (global stability) by resolving the ambiguity, either gradually or instantly. Finally, if listeners already know the song or the musical style of the song, they may avoid the garden-pathing altogether and tap in duple meter from the very beginning.

Method

Participants: Twenty subjects divided into two groups participated in the study (mean age = 26.5). The extent of their prior musical training ranged from none (3), some in the past (10), up to active involvement as music students or musicians (7). From post-experimental questioning, 12 participants reported never hearing the song before and the remaining eight reported knowing the song from somewhat to fairly well. The practice trials (utilizing various pieces in both duple and triple meter) ensured that the participants were able to synchronize their tapping with the music and that they could distinguish between duple and triple meters.

Equipment: Participants tapped on a Casio KW 1525 keyboard. Both stimulus presentation and data collection were controlled by and programmed in Max 4.0, run on an Apple G4 Powerbook. Music was played back through high-quality Edirol speakers.

Design: Participants first heard the whole song; in subsequent trials they were presented only with the target segment from the beginning of the song (80 sec long) consisting of four sections in which the ambiguity is most clearly present. Metrically indeterminate distractors (piano music by Schoenberg) were presented between the trials to prevent or minimize carryover effects.

Task: **The familiarization trial (Trial 1).** For the whole song presentation all participants were instructed to tap the meter (pattern of beats) that came most naturally to them (which they spontaneously perceived). The reason for using the same task for both groups in Trial 1 was to verify that there were not any substantial differences in performance between the two groups prior to the experimental manipulation. Participants were asked to start tapping as soon as they understood where the beats should go. Even though all participants tapped at least till the middle of the song (for some 140 seconds), only the first 80 seconds (the target segment consisting of Sections 1-4) were used for the data analysis.

Trials 2-5 – repeated rehearing of the target segment. Starting with Trial 2 the two groups were given different instructions. Participants assigned to the natural condition received the same instructions as in Trial 1, that is, to tap whatever felt most natural to them. In the coherent condition, participants were asked to come up with a coherent solution for tapping the beats over the course of the 80 second long segment. Before Trial 2 they were explicitly told that the meter of the song is ambiguous and that their task in the remaining trials was to come up with a single coherent way of tapping throughout the segment they would hear. They were not informed what the coherent tapping was meant to be.

Procedure: Participants were instructed to tap on two adjacent white keys (marked with different colors) on an electronic Casio keyboard, which produced two different snare-drum sounds (low and high). The lower key was reserved for tapping the downbeats and the upper key for tapping the remaining beats in a measure. Thus, to tap in triple meter a participant would tap ‘Lower-Upper-Upper’ (L-U-U, L-U-U, etc...). To tap in duple meter, a participant would alternate the two keys (L-U-L-U, L-U-L-U, etc...).

Criteria for evaluating tapping responses. A tap was considered valid if it occurred within ±100 *ms* of the precise moment that either duple or triple beats would appear in the stimulus audio file, as measured with SoundEdit Pro software. In order to exclude randomly hit correct taps, there had to be three consecutive taps within a duple or triple meter to be counted as valid. Taps falling outside the 200 *ms* time window that was centered at the precisely measured beat positions were considered to be ametrical tapping or errors.

Each section of the song was analyzed separately from a triple and a duple perspective. A 100 percent duple response in one section would consist of 32 taps (which equals eight 4/4 bars or sixteen 2/4 bars); a 100 percent triple response amounted to 48 taps (corresponding to sixteen 3/4 bars). Along with the type of meter, we also analyzed the phase (the downbeat placement), that is, the temporal relationship between the taps on the lower and upper key and their positioning with respect to the sound stimulus (i.e., the beats in music). Tapping in either duple or triple meter as depicted in Figure 4 was evaluated as in-phase tapping. When the lower key was pressed at beat locations other than the duple or triple downbeats, such tapping was evaluated as out-of-phase or shifted tapping. In the Results section below, we will discuss two sorts of analyses: one examines only in-phase tapping and the other combines both in-phase and out-of-phase tapping (which we refer to as *combined* tapping analysis).

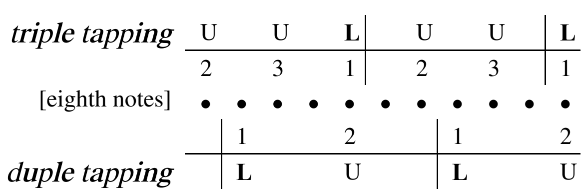


Fig. **4**. The lower (**L**) of the two neighboring white keyboard keys was used to indicate the perceived downbeat; the upper (U) key was reserved for weak beats. The relationship between the duple and triple downbeats was antiphasic.

Results

The tapping performance during the practice trials demonstrated that all the participants could synchronize their taps with the music and also that they were able to distinguish between duple and triple meters.

Familiarization trial (Trial 1). Because the instructions in this trial were the same in both conditions and because none of the tests comparing the two groups in each of the four sections of the target segment showed any significant differences, we plot the responses as one group. As Figure 5 shows, most participants began interpreting the piece in triple meter; as the piece continued a few began to doubt the triple interpretation and replaced it with a duple interpretation. Only two participants (n1 and c1) tapped duple meter exclusively despite the fact that they did not know the song beforehand. On the opposite extreme are two participants (c10 and n10) who tapped triple meter in every section (even though they eventually switched to duple tapping at some later point of the familiarization trial). While most of the participants replaced the triple mode with the duple mode at the beginning of the chorus, two participants (c2 and n2) made the switch early in the +Harmony section, presumably by redirecting their attention from the drum pattern to the pace of harmonic changes. During the first chorus, more participants tapped the duple meter and few continued in triple meter. Despite the fact that some later sections are virtually identical with the +Harmony section the vast majority of the participants continued tapping in duple meter and did not return to a triple interpretation. Looking at the averaged triple vs. duple tapping over four sections of the target segment, participants tapped more in triple than duple meter overall; this was true for the analysis of both in-phase tapping, $F(1,18) = 11.30$, $p < .005$, and combined tapping, $F(1,18) = 4.57$, $p < .05$.

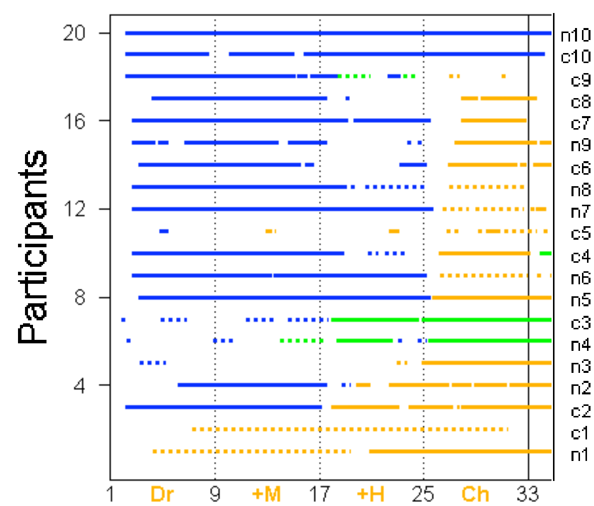


Fig. 5. Tapping in four sections (Dr, +M, +H, Ch - see Fig.3) during the familiarization trial. Participants from each group (n1-n10, c1-c10) are ordered based on the extent of their duple tapping. Orange lines = duple tapping, blue lines = triple tapping, green lines = '3over2' tapping[3], solid lines = in-phase, dotted lines = out-of-phase tapping, blank space = no or ametrical tapping.

Additional rehearings (Trials 2 through 5). With the next trial a series of four additional presentations of the first four sections of the song began. The main question was whether the participants, after hearing the whole song and after arriving at a duple interpretation during the initial hearing, would repeat the same process of 'being led down the garden path' or whether they would avoid garden pathing and start with the duple meter from the very beginning. Another question was how different instructions affected listening and tapping.

Trial 2. Figure 6 shows that participants in the natural condition were probably quite uncertain or even confused about how and what to tap, which resulted in a decrease of triple tapping without substantially changing the extent of duple tapping. In contrast, participants in the coherent condition were told before Trial 2 that the segment is metrically ambiguous. Figure 6 shows a marked change in proportion between duple and triple tapping for this group in Trial 2 (when compared with Trial 1). Repeated contrast tests conducted on each group separately showed a reliable 15 percent increase in duple combined scores in the coherent condition between Trial 1 and Trial 2, $F(1,9) = 6.25$, $p < .05$. However, because the tapping performance did not rise above 50 percent in one of the meters for the whole trial, listeners clearly did not im-

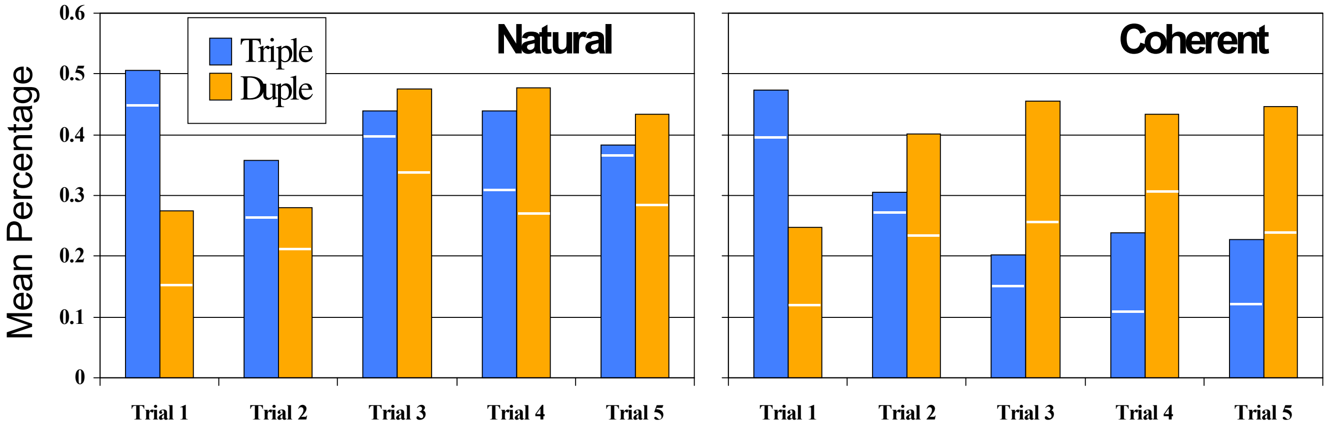


Fig. 6. The clustered bars show triple and duple tapping in each trial averaged over four sections. Each bar represents combined tapping; whereas the lower parts (below the white lines) show the extent of in-phase tapping, the upper parts show the out-of-phase tapping.

pose one coherent interpretation (with the exception of two subjects; see the accompanying file RMT-15Acc, Appendix 1, Trial 2, coherent condition).

Trial 3. With the second rehearing both groups started to show clear and stable differences. Participants in the natural group significantly increased their tapping scores in both meters as compared with the previous trial. The repeated contrasts analyses showed that in the natural group, Trials 2 and 3 were the only pair from all successive pairs of trials that showed reliable differences: $F(1,9) = 6.32$, $p < .05$ for duple combined tapping, and $F(1,9) = 8.90$, $p < .05$ for triple in-phase tapping. The increase of scores in both meters points to the fact that a majority of participants in the natural condition heard the stimulus as containing a change of meter. Six out of ten participants switched from triple to duple tapping right at the beginning of the +Harmony section. Such a uniformity in switching from one meter to another at the same location indicates that by the third encounter with the song these participants had built up expectations of an ostensible meter change after the +Melody section From looking at the graphs in Appendix 1 (see the accompanying file RMT-15Acc), it can be clearly seen how pervasive this interpretation was among the majority of the natural group in Trial 3 as well as in the remaining two trials.

Quite another picture emerges from looking at the development in the coherent condition. This group not only continued the tendency of increasing duple and decreasing triple tapping (the latter significantly so for triple in-phase scores, $F(1,9) = 5.28$, $p < .05$) but also reached the highest duple combined scores as well as the lowest triple combined scores in all five trials (Figure 6). Examination of individual performances in the coherent condition (see the accompanying file RMT-15Acc, Appendix 1, Trial 3) indicates that six out of ten participants avoided triple tapping completely. But obviously avoiding triple tapping in an attempt to come up with one coherent interpretation does not automatically or easily lead to imposing the duple meter. The following two trials clearly supported this observation.

Trials 4 & 5. As Fig. 6 shows, in the last two trials tapping performance did not change substantially in either group as compared with their performance in Trial 3. This observation was statistically confirmed by Helmert contrasts which examine the trends in the tapping responses over the course of the whole experiment[4]. The results of these analyses suggest that during the initial three presentations of the stimulus, participants have formed more or less stable representations (interpretations) of the music; there were no significant contrasts after Trial 2.

To sum up, the majority of participants in the natural condition seemed to arrive at the conclusion over the course of the experiment that the piece contains a change of meter. Thus the expectation that participants would progressively over multiple hearings replace triple tapping with duple tapping did not materialize in an overwhelming manner. Moreover, these data show that even when the presence of ambiguity is communicated to the listeners not all of them know how to go about resolving it. In fact, the majority of the coherent group (n = 6) did not resolve the ambiguity, even though half of them seemed to try to intentionally suppress the triple hearing. In both conditions in the familiarization trial (Trial 1), there was significantly more triple tapping in the target segment than duple tapping. Although in the following trials there was a tendency to increase the duple tapping in Trials 2 and 3, this tendency did not go beyond Trial 3; and also the triple scores did not decrease dramatically from Trial 2 on. With respect to the difference in instructions for the coherent group, the data show that it affected primarily the extent of triple tapping rather than duple tapping. Even though the triple scores in the coherent group were not significantly different from their duple scores, they were significantly lower from the triple tapping scores in the natural condition in Trials 3 to 5 (compare triple tapping between both groups in Fig. 6).

The individual tapping performances are summarized in Table 1. Clearly, the majority of listeners (70 percent; n = 14) did not apply duple tapping during the repeated presentation of the stimulus to any extent that would indicate coherent hearing of it, which allows us to conclude that these listeners processed the musical signal primarily in a bottom-up fashion. Thus, the notion of coherence of metrical structure does not seem to figure prominently in the awareness of these listeners. Out of the six remaining participants, half of them tapped exclusively in the duple meter. From this group two subjects reported good familiarity with the song (see the last line of Table 1; for this purpose we use Bharucha's [1] term 'veridical expectation'[5]). Only in three cases did participants recognize the garden path nature of the segment and reinter-

TABLE 1
Crosstabulation of the Number of Participants (in Each Group) with Types of Tapping Responses and Prediction Specification

Processing	**Natural** (n = 10)	**Coherent** (n = 10)	*Tapping response*	*Scenarios*
bottom-up	2	2	tapping triple or '3over2'	a
	6	1	switching from triple to duple	b
	-	3	inhibiting triple but trouble with duple	c
top-down	1	2	gradually imposing duple	c
	-	1	stylistically expecting duple	d
	1	1	veridically expecting duple	d

pret it by applying some kind of top-down listening strategy (e.g., inhibition and/or retrospection). Specific examples of individual differences that follow the structure of Table 1 can be found in Appendix 2 (see the accompanying file RMT-15Acc). Correlation analyses (after excluding one outlying case) showed an association between prior musical experience and the tapping performance ($r = .46$, $p < .05$) and no association between the extent of familiarity with the song and the tapping performance.

3. DISCUSSION

What are the possible explanations for this pattern of results? First, the ostensible change of meter apparently did not sound incoherent enough that these listeners felt prompted to do something about it. These listeners did not behave like the experienced listeners Lerdahl and Jackendoff [8] try to model, who—as they proposed—do more than merely perceiving sequences of notes; the experienced listeners are assumed to impose a deeper organization that makes the sequence *coherent*. Perhaps for some listeners music perception is more like a passive taking in of information from the environment than the more active structuring of information that Lerdahl and Jackendoff argue for.

Second, even among those participants who eventually arrived at a coherent (duple) hearing, none of them resolved the ambiguity instantly, that is, immediately after the first hearing. Thus, we found no evidence for the (more or less) instantaneous ambiguity resolution that the simultaneous presence of several analyses computed in parallel predicts [6]. Accordingly, one can argue that the process of ambiguity resolution in music takes place on a different time scale than in language, that is,—provided the ambiguity is detected—it more likely than not requires multiple rehearings. In sum, our data suggest that people's sensitivity to details and intricacies of musical (metrical) structure is of a different nature than to language structure (the recovery from garden path probably being only one example out of many).

From the perspective of modularity theory, this study raises the question whether higher-level functions (like memory and knowledge) may outweigh the output of an informationally encapsulated processor (the music parser). In order for this to happen, the initial bottom-up percept that has been led down the garden path has to be inhibited on subsequent hearing(s) after the contextually informed higher processing mechanism equipped with the previously absent knowledge voluntarily imposes top-down alternative listening strategies. As we have seen, this study provides a rather modest support for this possibility. But nevertheless it shows that our perception of a piece may change over the course of multiple rehearings or, alternatively, may be affected by our long-term knowledge of it or by our acquaintance with a style to which the piece belongs.

But, apparently, for a majority of listeners the parser continues to process the piece as if it were being heard for the first time; as Jackendoff put it, the parser generates expectations that we know will be denied and wanders naively down garden paths. This study makes a strong case for such a modular account.

Alternatively, one could argue that the particular song used as the stimulus in this experiment is an extreme case in which the insistent cross-accentuation (i.e., syncopation) throws the listener's sense of meter off. It is possible that the lack of any prior duple context as well as the sparsity of acoustic cues for the duple meter in the drum introduction cements listeners conviction that the song starts out in triple meter regardless of what is happening later on. As a result, the listeners may not even consider reinterpreting the drums within the duple framework.

To test these alternatives, the next step in this project is to extend the number of conditions via introducing contextual priming that might facilitate the reinterpretation of the drum section. In the original version, each of the sections following the drum introduction contains progressively stronger disambiguating cues. The contextual manipulations will reorder the sections so that the section that was originally positioned at the very beginning (the Drum section) is preceded by a section or sections that originally appeared later.

If the garden-path effect is attenuated or even eradicated as a result of this contextual manipulations, this will show that immediately preceding metrical context can affect the interpretation of an ambiguous passage. If this kind of contextual effect extends to subsequent rehearings of the original song, this could suggest how top-down listening strategy that leads to avoiding the garden path might be induced.

Notes:

[1] On the other hand, even though Lerdahl admits that some listeners may hear global structures differently, he nevertheless argues that listeners, in general, hear pieces as ending functionally on the tonic and thus perceive tonal closure regardless of undetected key changes.

[2] Here we have allowed ourselves one methodological simplification by treating one 4/4 (or 12/8) measure of the song as consisting of two 2/4 (or 6/8) measures. In doing so, we have not sacrificed much from the underlying structural complexity; rather, this simplification makes the comparison between duple and triple meter as it pertains to the present stimulus more straightforward.

[3] '3over2' mode of tapping is an inherently triple mode, but one whose phase relationship follows duple cues (the chord changes). As a result, this tapping mode is antiphasic to the initial triple meter. We cannot explain this in a more detail here but overall this mode of tapping amounts only to 4-6 percent of all produced tapping in each group.

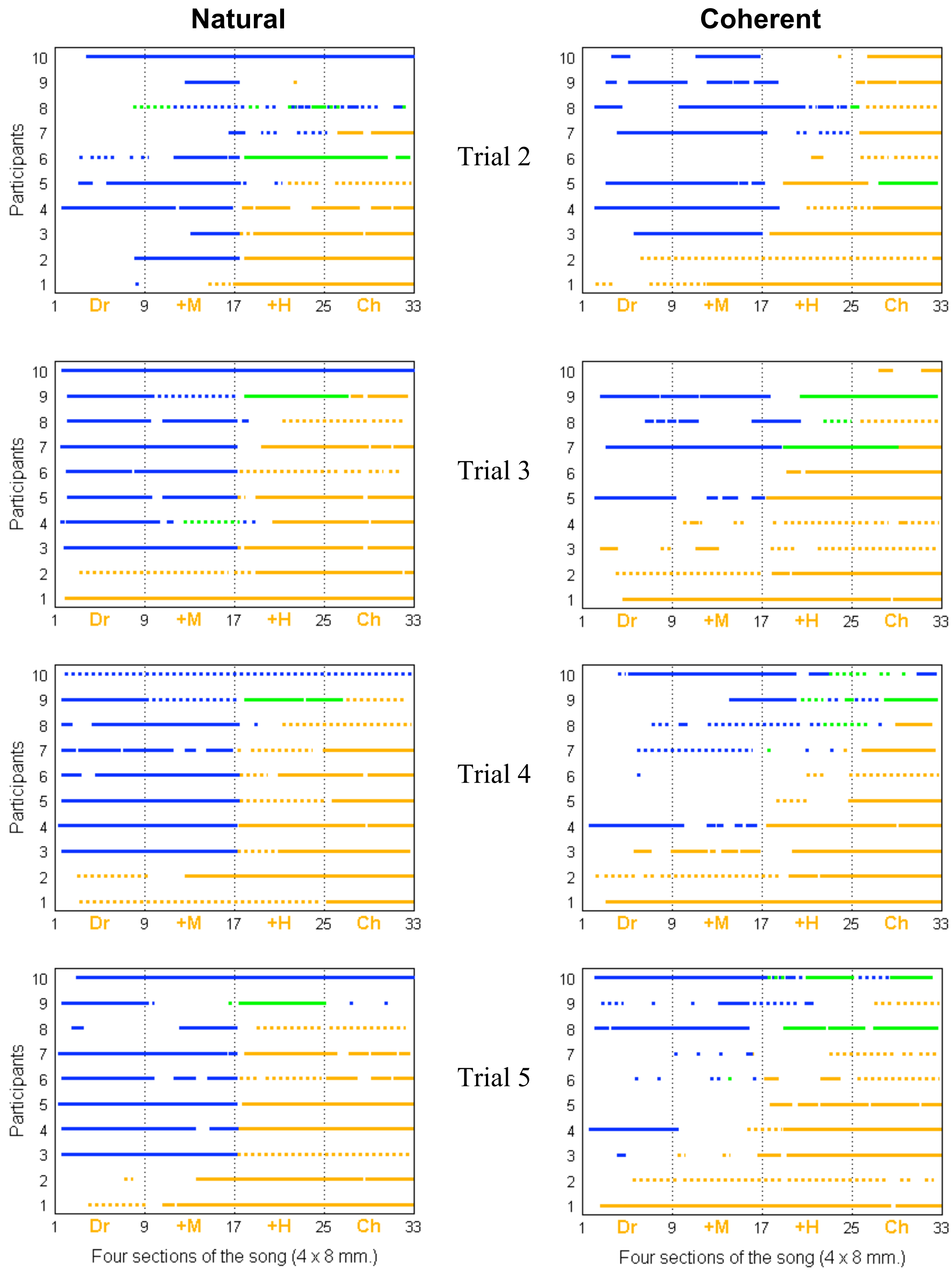
[4] These tests compare in succession each trial with the mean of all subsequent trials, a procedure that enabled us to determine whether there was a point in the experiment after which no substantial changes in the tapping performance occurred.

[5] However, the utility of the term 'veridical expectation' is somewhat questionable in the present context because these two subjects (see graphs (*g*) and (*h*) in Appendix 2 located in the accompanying file RMT-15Acc) despite having tapped exclusively in duple meter have rarely or not at all followed the 'veridical' downbeat placement. That is, they mostly tapped out of phase.
Moreover, one can argue that the memory that this piece contains a change of meter (even if it really does not) is equally veridical for those who remember the piece that way. Thus, the meaning of 'veridical' seems to come in two flavors. Whereas the weak one holds that veridical is whatever one believes as being true, the strong version sticks with what can be established as objective reality. For instance, Oxford dictionary provides only the latter version (veridical = truthful, coinciding with reality).

4. REFERENCES

1. Bharucha, J. J. (1994). Tonality and expectations. In R. Aiello & J. A. Sloboda (Eds.), *Musical perceptions* (pp. 213-39). New York: Oxford University Press.
2. Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception, 5*, 197-205.
3. Handel, S. (1998). The interplay between metric and figural rhythmic organization. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1546-1561.
4. Hodges, D. & Lipscomb, S. (2001). *Tonality judgments in popular music contexts by preteens and college students: A comparative analysis.* Unpublished manuscript.
5. Jackendoff, R. S. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.
6. Jackendoff, R. S. (1991). Musical parsing and musical affect. *Music Perception, 9*, 199-230.
7. Jackendoff, R. S. (1992). *Languages of mind: Essays on mental representation.* Cambridge, MA: MIT Press.
8. Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.
9. Marvin, E. W. & Brinkman, A. (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception, 16*, 389-407.
10. Palmer, C. & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16,* No.4, 728-741.
11. Povel, D. J. & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*, 411-440.
12. Snyder, J. & Krumhansl, C. L. (2001). Tapping to Ragtime: Cues to Pulse Finding. *Music Perception, 18*, 455-489.
13. Sting & Summers, A. (1996). Murder by Numbers. In The Police, *Message in a box. The complete transcriptions*, *Vol. 4*, (p. 46), Transcribed by A. Robyns. Milwaukee, WI: Hal Leonard.
14. Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 18, 645-659.
15. Swinney, D. (1982). The structure and time-course of information interaction during speech comprehension: Lexical segmentation, access, and interpretation. In J. Mehler, E. Walker, and M. Garrett (Eds.), Perspectives on mental representation. Hillsdale, NJ: Erlbaum.
16. Tanenhaus, M., Leiman, J. & Seidenberg, M. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. Journal of Verbal Learning and Verbal *Behavior, 18,* 417-440.
17. Vazan, P. & Schober, M. F. (2002). [Influence of context on meter disambiguation]. Unpublished raw data.
18. Yee, W., Holleran, S., & Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: Influences of musical skill. *Perception & Psychophysics, 56*, 461-471.

Appendix 1:



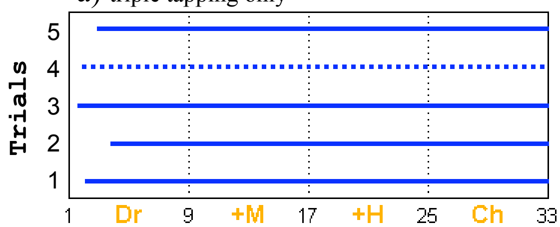
Tapping in four sections (Dr, +M, +H, Ch - see Fig.3) of the target segment in Trials 2 to 5. Participants from each group are ordered based on the extent of their duple tapping. Orange lines = duple tapping, blue lines = triple tapping, green lines = ‘3over2’ tapping (see Note 3), solid lines = in-phase, dotted lines = out-of-phase tapping, blank space = no or ametrical tapping.

Appendix 2:

Examples of tapping response types from Table 1

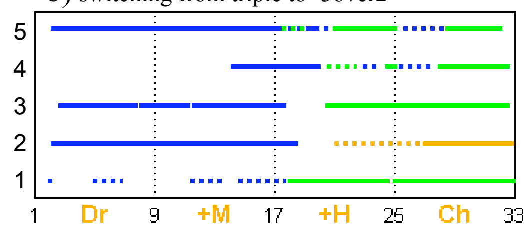
The following graphs show eight individual tapping responses (three from the natural condition and five from the coherent condition); each line represents separate trial (see the explanation of colors and line types below)

a) triple tapping only



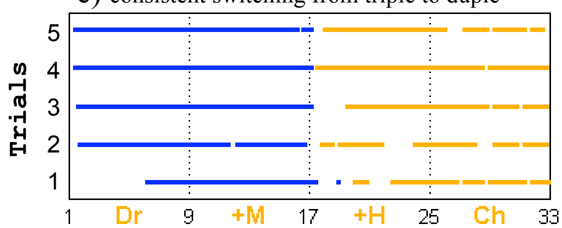
the only instance of tapping consistently in triple meter in every trial

b) switching from triple to ‘3over2’



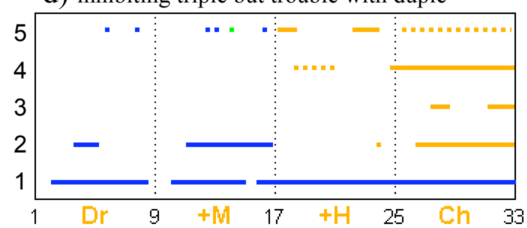
one of three examples of switching from triple to ‘3over2’ tapping (see Note 3)

c) consistent switching from triple to duple



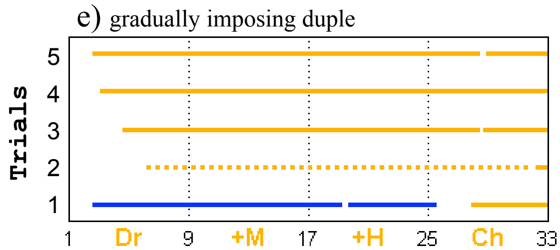
most frequent response (especially in the natural condition) indicating a perception of changing meters at the beginning of +Harmony section

d) inhibiting triple but trouble with duple



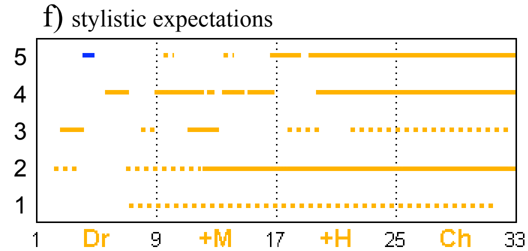
one of three instances of inhibition of triple tapping in the coherent condition (note that there was no duple tapping instead in the first two sections)

e) gradually imposing duple



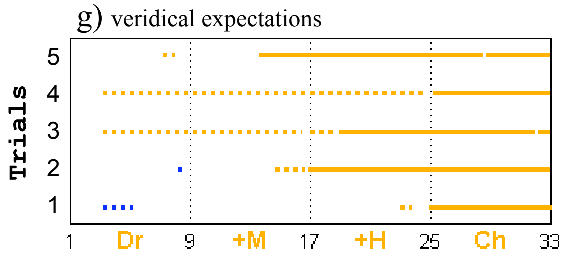
paradigmatic example of gradual ambiguity re-solution—note that the tactus and the downbeat placement (the phase) were changed at different times

f) stylistic expectations

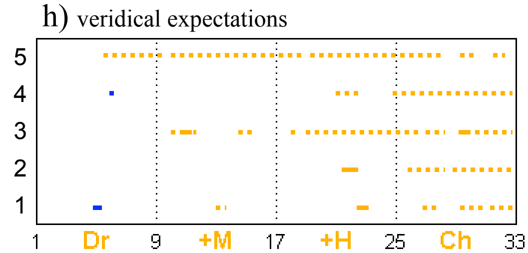


the participant begins tapping in duple meter even though he did not know the song already; perhaps his familiarity with styles like ‘reggae’ or ‘ska’ helped him infer the duple structure right from Trial 1

g) veridical expectations



h) veridical expectations



the only two cases where almost no triple tapping was observed and where familiarity with the song was also reported —however, note the absence of in-phase tapping during the first two sections in g) as well as its complete absence in all sections in h) (in out-of-phase tapping it is most often the bass drum that is identified as the downbeat)

Tapping in four sections of the target segment over five trials. Orange lines = duple tapping, blue lines = triple tapping, green lines = ‘3over2’ tapping, solid lines = in-phase tapping, dotted lines = out-of-phase tapping, blank space = no or ametrical tapping.