Journal of Experimental Psychology:
Human Perception and Performance
1998, Vol. 24, No. 1, 252–260

0096-1523/98/$3.00

Implicit Knowledge Versus Psychoacoustic Similarity in Priming of Chords

Hasan Gürkan Tekman
Middle East Technical University

Jamshed J. Bharucha
Dartmouth College

A chord-priming paradigm was used to test predictions of a neural net model (MUSACT). The model makes a nonintuitive prediction: Following a prime chord, expectations for the target chord are based on psychoacoustic similarity at short stimulus onset asynchronies (SOAs) but on implicit knowledge of conventional relationships at longer SOAs. In a critical test, 2 targets were selected for each prime. One was more psychoacoustically similar to the prime, and the other was more closely related on the basis of convention. With an SOA of 50 ms, priming favored the psychoacoustically similar target; with SOAs of 500 ms and longer, the effect reversed, and priming favored conventional relatedness. The results underscore the limitations of models of harmony based on psychoacoustic factors alone. These studies demonstrate how neural net learning models that are appropriately constrained can be subject to strong empirical verification.

It is an age-old question whether relationships between musical chords are driven by the physical structure of periodic sounds or by implicit knowledge of cultural convention. In advancing the former view, Helmholtz (1885/1954) suggested a measure of acoustic similarity based on the number of frequencies in common between two chords. A contemporary descendent of this approach (Parncutt, 1989; Terhardt, 1974, 1979) uses a measure of psychoacoustic similarity based on the number of evoked pitches in common. The implicit knowledge hypothesis receives support from the priming of chords, which seems to be based on conventional associations rather than shared frequencies (Bharucha & Stoeckig, 1987). In this article, we explore some of the boundary conditions of these two views.

Priming tasks reveal that people have substantial implicit knowledge of conventional relationships between chords, even if they have had no formal musical training and cannot verbalize this knowledge (Bharucha & Stoeckig, 1986, 1987). In particular, for Western participants with or without formal musical training, a chord primes other chords that have high transition probabilities in the most prevalent (i.e., popular) forms of Western music. These priming effects have been simulated by a neural net model (MUSACT) that has passively internalized the grouping of tones typical of the Western musical environment (Bharucha, 1987a, 1987b). It has been suggested that pervasive tonal patterns become encoded through neural self-organization (Bharucha, 1991; Gjerdingen, 1990). The MUSACT model represents the end-state of such a process as applied to the most typical grouping patterns of tones and chords. It is thus a model of the implicit knowledge acquired through extended exposure to the conventional relationships of a musical culture. The more likely two chords are to occur in the same piece or segment thereof, the greater their conventional relatedness and the stronger their association in the model.

We refer to tones, major chords, and major keys by using letter names with subscripts, $_t$, $_c$, or $_k$, respectively. For example, the set of tones (G_t, B_t, D_t) constitutes the G major chord, G_c, and the set of chords (C_c, G_c, D_c) evokes the G major key, G_k. For the most part, the more conventionally related two chords are, the greater their psychoacoustic similarity. For example, C_c is conventionally more closely related to G_c than to D_c because C_c occurs in two keys with G_c—the keys of G_k (C_c, G_c, D_c) and C_k (F_c, C_c, G_c)—but in only one key with D_c—the key of G_k (C_c, G_c, D_c). C_c (C_t, E_t, G_t) is also psychoacoustically more similar to G_c (G_t, B_t, D_t) than to D_c (D_t, $F\#_t$, A_t) because C_c shares one component tone (G_t) with G_c, but it shares none with D_c. The close correspondence between conventional relatedness and psychoacoustic similarity suggests that the latter may have influenced the development of the former.

Parsimony dictates that if psychoacoustic similarity is sufficient to account for relationships in musical harmony, the postulation of implicit knowledge is unnecessary. In this article, we consider a set of cases in which the two hypotheses make opposite predictions. If the data favor psychoacoustic similarity in these divergent cases, then the postulation of implicit knowledge—and a model thereof—would be unnecessary. If not, psychoacoustic similarity would be shown to be insufficient as an account of musical harmony, and implicit knowledge would be implicated.

Although C_c and D_c occur together in G_k (C_c, G_c, D_c), C_c and E_c do not occur together in any key. Consequently, if priming is driven by implicit knowledge of conventional

---

Hasan Gürkan Tekman, Department of Psychology, Middle East Technical University, Ankara, Turkey; Jamshed J. Bharucha, Department of Psychology, Dartmouth College.

This work was completed while Jamshed J. Bharucha was a Fellow at the Center for Advanced Study in the Behavioral Sciences. Research was supported by National Science Foundation Grants DBS-9222358 and SES-9022192.

Correspondence concerning this article should be addressed to Jamshed J. Bharucha, Department of Psychology, 6207 Gerry Hall, Dartmouth College, Hanover, New Hampshire 03755. Electronic mail may be sent via Internet to bharucha@dartmouth.edu.

relationships, C_c should prime D_c more strongly than it should prime E_c. If, on the other hand, priming is driven by psychoacoustic similarity, C_c should prime E_c more strongly than it should prime D_c. This is because C_c (C_t, E_t, G_t) shares a component tone (E_t) with E_c (E_t, $G\#_t$, B_t) but not with D_c (D_t, $F\#_t$, A_t).

Psychoacoustic Similarity

A contemporary measure of psychoacoustic similarity derives from Terhardt's (1974, 1979) theory of pitch and its recent elaboration and modification by Parncutt (1989). According to Terhardt, a spectrum can evoke a multiplicity of pitches (i.e., pitch cues). Spectral pitch cues correspond to frequencies present in the signal (perhaps slightly shifted because of masking). Virtual pitch cues are evoked at the fundamental frequencies of subsets of the spectral cues. For example, when the fundamental frequency of a harmonic spectrum (such as a voiced speech sound or a musical instrument tone) is missing, one nevertheless hears a pitch corresponding to that frequency. Parncutt suggested a measure of psychoacoustic similarity, called pitch commonality, based on the correlation between the pitch cues evoked by two signals. For purposes of this article, pitch commonality can be taken to be the measure of psychoacoustic similarity.

Chord Priming

Priming enables us to measure the automatic expectations induced by a musical context. Bharucha and Stoeckig (1986, Experiments 2 and 3) presented participants with two chords in succession, a prime followed by a target. On half of the trials, the target was mistuned slightly to create a foil. Participants were asked to identify the second chord as a target or a foil by making an intonation judgment, that is, to indicate whether the second chord was in-tune or out-of-tune. The prime and target varied in the degree of conventional relatedness. We refer to the more closely related pair as *close* and the more distantly related pair as *distant.* Even though the close and distant pairs were equally likely to occur in the experiment, participants were faster and more accurate for the close pairs. We refer to the response time and accuracy advantage for close targets over distant targets as the *chord-priming effect.*

For the foils, the effect was reversed, though of smaller magnitude, as reflected by an interaction between relatedness and intonation. In other words, when the two chords were closely related, the second chord tended to sound in-tune; when the two chords were distantly related, the second chord tended to sound out-of-tune. The prime's effect was not only to facilitate the intonation decision of related successors but also to bias their intonation so that the more distant the relationship the greater the dissonance. Responses to the foils thus reveal that priming takes the form of an overall processing advantage as well as a perceptual bias.

What accounts for the effect of a chord on the perception of what follows? Because the related pairs in the Bharucha & Stoeckig (1986) experiments also had more harmonics in common, the observed effect could have been due to psychoacoustic similarity (in this case, repetition priming at the level of shared frequencies). However, priming persisted when shared harmonics were removed, suggesting that it must be a consequence of conventional associations encoded at a cognitive level and cannot be explained solely by shared frequencies (Bharucha & Stoeckig, 1987).

Parncutt (1989) has since argued persuasively that although spectral pitch cues shared by prime and target were eliminated in the Bharucha and Stoeckig (1987) study, the missing harmonics may have been restored as virtual pitch cues—pitch sensations corresponding to frequencies that are not present in the signal—as predicted by Terhardt's (1974, 1979) model of pitch. In the present article, we seek to overcome this confound by using prime–target pairs in which the close pair has no advantage over the distant pair in terms of either spectral or virtual pitch cues in common.

In order to conduct the critical test, we exploited the existence of chord pairs for which psychoacoustic similarity and implicit knowledge of conventional relatedness make opposite predictions. For example, C_c is more closely related to D_c than to E_c but more psychoacoustically similar to E_c than to D_c. C_c shares a key with D_c, but it shares a component tone (E_t) with E_c. Chords that share a component tone are more psychoacoustically similar than chords that do not (they also share the harmonics and subharmonics of this tone). Krumhansl, Bharucha, and Kessler (1982) found that participants judge C_c and D_c to be more closely related than C_c and E_c, so there is already evidence from rating judgments to reject the sufficiency of psychoacoustic similarity. The MUSACT model not only extends the relatedness prediction to priming but also generates a strong nonintuitive prediction: Although relatedness should drive priming when the network has settled (i.e., when the constraints implicit in the connectivity have been satisfied and the qualitative pattern of activation no longer changes), shared tones should drive priming in the earliest activation cycles. In other words, priming should favor E_c at arbitrarily short stimulus onset asynchronies (SOAs) and favor D_c thereafter.

The MUSACT Model

The pattern of clustering of tones is among the persistent regularities in musical forms that are pervasive in the West. Highly constrained combinations of tones are used simultaneously or in close succession to form explicit or implied chords. These chords, in turn, are used in highly constrained combinations, which suggests familiar abstract units called *keys.* Although this pattern of clustering by no means captures the structure of music in all its complexity and subtlety, and it may not characterize many genres, it is nevertheless one of the most predictable properties of the music that is most widely heard in the West, as evidenced by the song books that provide the chords for popular songs.

The MUSACT model consists of a network that maps tones to chords and that maps chords to keys, as shown in Figure 1 (from Bharucha, 1987a). Units in the tone layer are assumed to be tuned to pitch classes (i.e., tones generalized across octaves). Although only 12 tones corresponding to the chromatic scale are shown, this layer is presumed to be

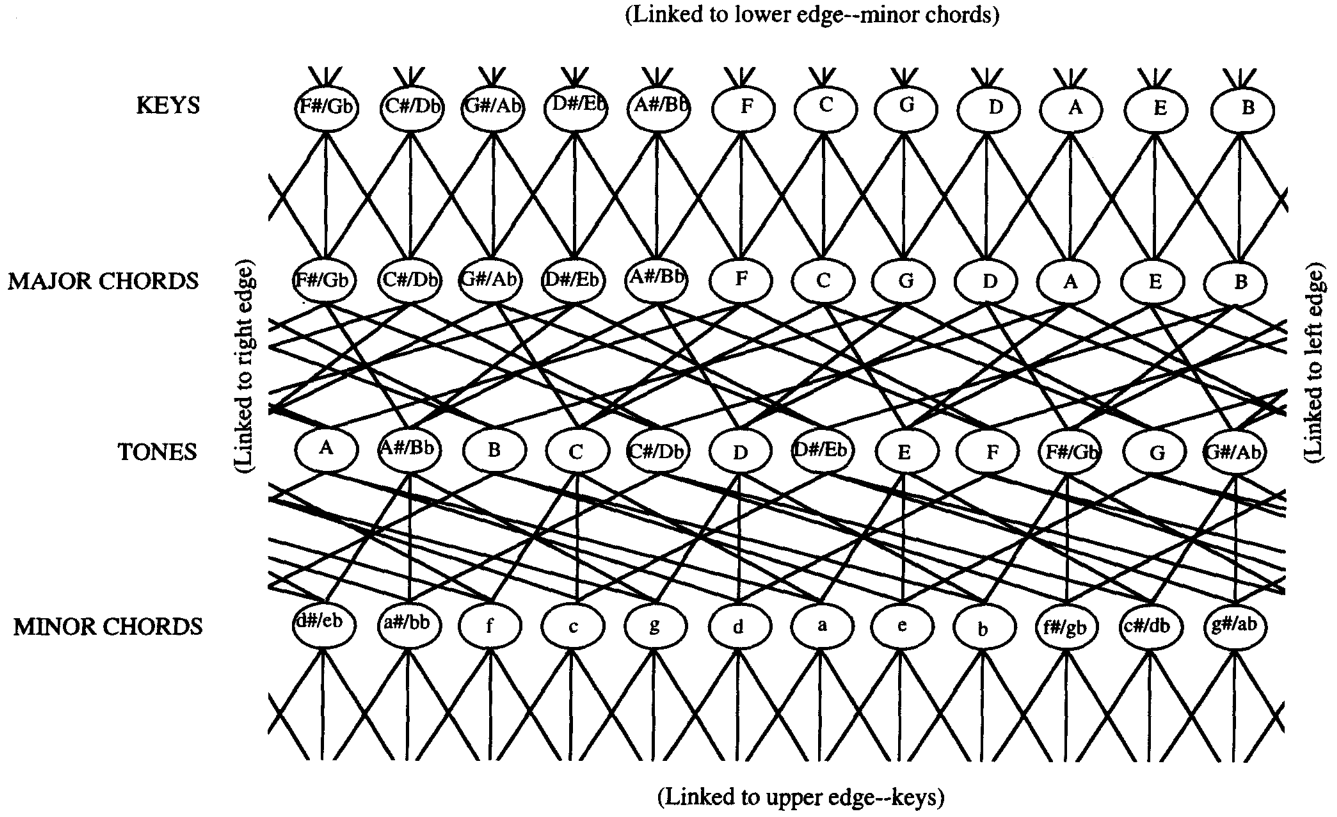

Figure 1. The MUSACT network. From "MUSACT: A Connectionist Model of Musical Harmony" (p. 511) by J. J. Bharucha, 1987a, in *Ninth Annual Conference of the Cognitive Science Society,* Hillsdale, NJ: Erlbaum. Copyright 1987 by the Cognitive Science Society. Reprinted with permission.

much denser. Units in the chord layers are indirectly tuned to chords by virtue of their connectivity with the tone units. (Units for major and minor chords are shown on different sides of the tone units for graphic convenience only.) Similarly, units in the key layer are indirectly tuned to keys by virtue of their connectivity with the chord units. In order to visualize the structure, Figure 1 depicts only the links that have high connection weights: Each chord unit is actually linked to each tone unit, and each key unit is actually linked to each chord unit. The indirect tuning of each chord or key unit is determined by the distribution of connection weights on the links that feed into it. These weight distributions—and consequently, these indirect tuning characteristics—are learned by passive self-organization (Bharucha, 1991).

Through self-organization, a major chord—perhaps the most pervasive tonal pattern in Western music—becomes encoded in the connection weights that feed into a single category unit from units tuned to the component tones of the chord. The consequence of this form of learning is the development of category units that serve as feature detectors for typical chords (see Figure 1). Analogously, chords that are typically grouped together in a piece of music or segment thereof and that evoke a sense of key drive the self-organization from the chord units to a more abstract layer of category units that become feature detectors for keys.

The unit in the network that comes to be the C_c unit does so because of the strengthening of links from the C_t, E_t, and G_t units. (Figure 1 depicts only these strong links.) Analogously, the unit in the network that comes to be the C_k unit does so because of the strengthening of links from the F_c, C_c, and G_c units. Each tone is a component of three different major chords, and each chord is a member of three different major keys. The connectivity in Figure 1 reflects these relationships. The network thus represents the implicit schematic knowledge a person develops about conventional groupings of tones and chords as a result of extended exposure to these regularities.

Algorithms for neural self-organization (e.g., Fukushima, 1975; Grossberg, 1970, 1976; Kohonen, 1984; Rumelhart & Zipser, 1985; von der Malsberg, 1973) can account for the formation of hierarchical encodings in which events that often occur together are clustered together. Although the algorithms differ in how they handle the resolution and stability of clusters, they are grounded in the same underlying learning principles. The weights on the connections between a chord unit and the 12 tone units are a vector in 12-dimensional tone space. Each chord unit can thus be thought of as a weight vector. The pattern of activation across the tone units in response to a cluster of tones is also a vector in this space. A chord unit is tuned to a particular cluster of tones when its weight vector is collinear with the activation vector of tones that make up that chord. Self-organization tends to move weight vectors toward activation vectors or principle components thereof. Hence, the idealized connectivity of the MUSACT network reflects the end-state of self-organization in response to a musical environment in which major and minor chords are typical tone clusters. The activation vectors are assumed to be decaying memories or temporal composites, that is, they integrate activation over time (Bharucha, in press). Thus, typical sequences of chords would produce vectors that are sufficiently similar across temporal orders to be self-organized as keys. Musical environments with different clustering regularities would of course yield different patterns of connectivity, and the units in the self-organized layer would be tuned to structures other than the chords and keys of this network. The network is thus the internalization of some clustering regularities of a musical culture.

The predictions tested in the following experiments are based on the end-state of the network. When a chord is presented to the network, the tone units tuned to the sounded tones are activated. For example, if C_c is sounded, tone units C_t, E_t, and G_t are activated. The spread of activation is phasic, that is, a unit's activation changes as a function of the change in activation of another unit connected to it. Connections are bidirectional. On any given update cycle, the activation of each unit is its activation from the previous cycle plus a weighted sum of the phasic activations received from connected units. This process continues until a state of equilibrium is reached, in which activation levels of all units do not change more than a criterion amount on successive activation cycles (see Bharucha, 1987a).

The extent to which a potential target chord is primed at any given time following the presentation of the prime is a function of the activation of that target unit at that time. Just how activation influences a participant's decision is presumed to be determined by a decision system that receives input from the network and is outside the scope of the MUSACT model. Whatever the precise decision system, one must account for how a prime chord privileges some potential targets over others. We take activation to be the underlying privileging process that informs the decision.

When the system reaches equilibrium, the relative activations of chord units reflect their degree of relatedness to the sounded chord. Activation of the chord units decreases as the distance from the stimulus chord in the circle of fifths increases (Figure 2B). Note that nothing about the circle of fifths was explicitly built into the network; it is, in this sense, an emergent property of the learned connectivity. (The spatial layout of chord units in order of the circle of fifths in Figure 1 is for visual convenience only, and it reflects the minimally tangled configuration.) The circle of fifths can thus be taken to be a metric of conventional relatedness. Although conventional relatedness (distance around the circle) and psychoacoustic similarity are correlated, there are cases in which one pair of chords is more psychoacoustically similar whereas another is more closely related. These cases provide a critical test because the MUSACT model and models based on psychoacoustic similarity make opposite predictions.

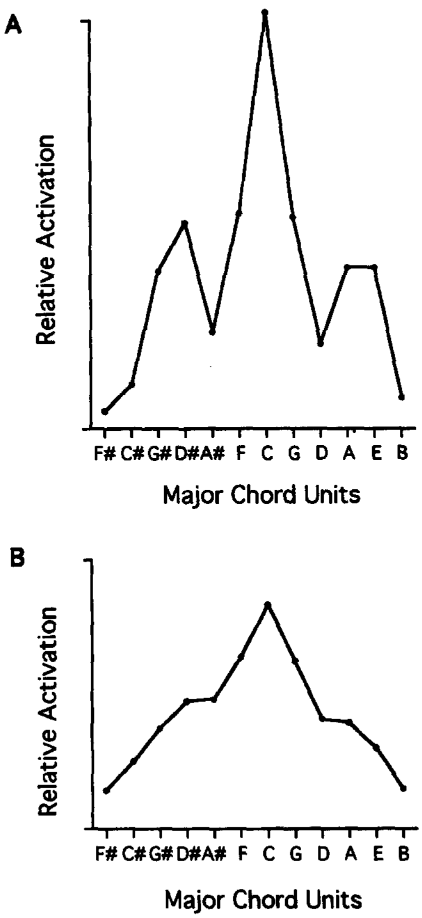

Figure 2. Relative activation of major chord units in the network following the presentation of a C major chord. (A) At an early stage of processing, activation is based on common tones. (B) After the network reaches equilibrium, activation is based on conventional relatedness. Based on a computer simulation of the original model used in "Music Cognition and Perceptual Facilitation: A Connectionist Framework" by J. J. Bharucha, 1987b, *Music Perception, 5,* pp. 1–30. Copyright 1987 by the University of California. Adapted with permission.

A nonintuitive prediction of the model was made in an earlier article (Bharucha, 1987b) and is tested here: The pattern of activation changes qualitatively as the network settles, such that in early activation cycles, the pattern reflects the number of tones shared by prime and target, whereas at later activation cycles, the pattern reflects conventional relatedness. This is because a sounded chord initially activates all and only chords that have tones in common with the sounded chord. At this early stage of activation, the chord units are divided into three groups with regard to their activation (Figure 2A). The unit representing the sounded chord (for illustration, C_c) has the highest activation because it receives activation from all three tone units to which it is connected. Next most active are units representing major chords that share one tone with the sounded chord (chord units $G\#_c$, $D\#_c$, F_c, G_c, A_c, E_c). The remaining major chord units have no activation because they share no tones with the sounded chord.

When activation is allowed to reverberate through the network, flowing from chord units to key units and back to chord units, the pattern of activation at the chord units is altered qualitatively and begins to reflect conventional relatedness rather than common tones (Figure 2B). The dynamics of the activation process of selected units are shown in Figure 3. G_c and E_c share a tone with the sounded chord, C_c, and therefore have high relative activation in the very first activation cycle. D_c and $F\#_c$ share no tones with the sounded chord and receive little activation at this stage. As activation spreads, the qualitative pattern changes to reflect conventional relatedness. In particular, D_c and E_c cross over. The ensuing empirical prediction is that priming should be driven by the psychoacoustic similarity engendered by common tones at short SOAs and by conventional relatedness at long SOAs.

Experiment 1

In this experiment, the prime and target are two steps apart on the circle of fifths in one condition (close condition) and four steps in the other condition (distant condition). If the prime is C, the close target is D and the distant target is E in this experiment. The close target shares a key cluster with the prime, whereas the distant target does not. However, the distant target shares a tone with the prime, whereas the close target does not. This pits conventional relatedness against psychoacoustic similarity. The model predicts stronger priming for close than for distant targets, with a reversal at very short SOAs.

Method

Participants. Twenty-one Dartmouth College undergraduates taking an introductory psychology course participated in the experiment. All participants reported having normal hearing. The mean number of years they had played a musical instrument or performed with voice was 4.55. The range of experience with musical instruments or voice was 0 to 14 years.

Apparatus. Stimuli were created by an Apple Macintosh II computer. A Sansui A-707 amplifier and Sennheiser HD-410 headphones were used for stimulus presentation.

Stimuli. Stimuli were major chords synthesized with equal tempered tuning (in which one semitone corresponds to a frequency factor of $2^{1/12}$). A major chord consists of three component chromas: the *root*, the *third* (Root × $2^{4/12}$), and the *fifth* (Root × $2^{7/12}$). Each of the three chromas was represented by a tone in each of five octaves. Each component tone was a complex tone that included the first four harmonics at equal amplitudes. Loudness of the tones tapered off to the hearing threshold in the outer octaves (after Krumhansl et al., 1982; Shepard, 1964). This method maximizes the sense of chroma and minimizes the effect of pitch height. The root tone of the prime chord was chosen at random from the 12 chromatic tones.

The out-of-tune chords (foils) were created by reducing the frequency of the fifth in the major chord by a frequency factor of $2^{3/96}$.

Procedure. Participants were given a training session to ensure that they could discriminate between targets and foils at a criterion level of accuracy. Participants initiated each trial with a button press. In each training trial, a chord was played for 2 s (with no decay), and participants were required to press one of two buttons labeled *in-tune* and *out-of-tune* on a response box. Feedback about accuracy was given after each trial. The training session consisted of 48 trials, and the criterion number of correct responses was 43 (approximately 90% of the trials). Participants who performed below the criterion were given a choice of repeating the training or withdrawing from the experiment. Only participants who performed at or above the criterion participated in the actual experiment.

In each trial of the main experiment, a button press initiated the trial. Each trial began with a sequence of 16 tones whose pitches were chosen at random. All tones were of equal duration, and the entire sequence lasted 2 s. This was intended to erase possible effects of memory for the chords from the preceding trial. After a

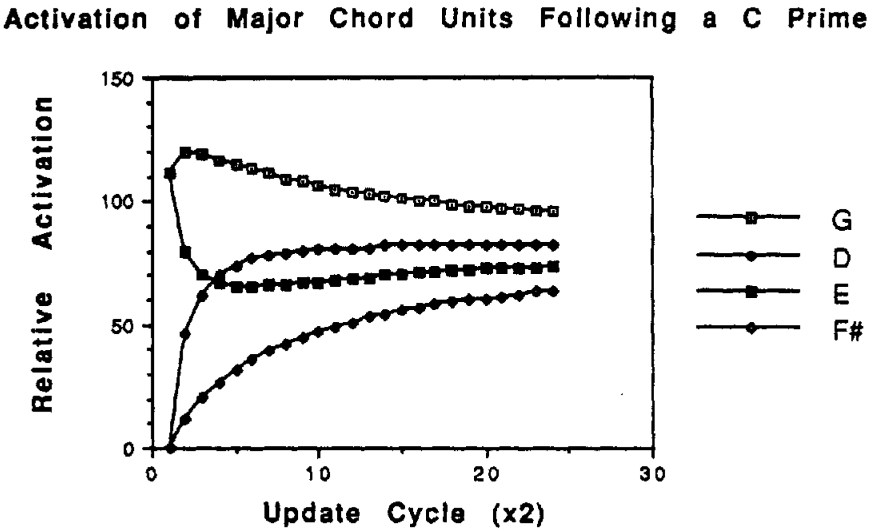

Figure 3. Relative activation of major chord units over time, following the presentation of a C major chord.

short silence of variable duration, two chords were played without any break in between. Participants were instructed to decide whether the second chord was in-tune or out-of-tune and to respond with a button press as in the training. Feedback on the accuracy of the response was given after each trial.

The four different durations of the prime chord were 50, 500, 1,500, and 2,500 ms. The time from the onset of the random sequence to the onset of the second chord was held constant at 6 s. The second chord was always played for 2 s. Participants were instructed to respond during the 2 s that the chord was played.

In the close trials, the root tones of the two chords were separated by two semitones (a frequency factor of $2^{2/12}$); in the distant trials, the roots of the two chords were separated by four semitones (a frequency factor of $2^{4/12}$).

The design was completely crossed, with two levels of relatedness (close vs. distant), two levels of intonation (in-tune vs. out-of-tune), and four prime durations (50, 500, 1,500, or 2,500 ms). There were 14 replications of each of the 16 conditions, yielding 224 trials. On each trial, the root tone of the prime chord was chosen at random from the 12 chromatic tones. Trials were blocked by the duration of the prime. The order of blocks was chosen at random for each participant.

In order to familiarize the participants with the task, a sequence of 16 practice trials preceded the experiment. The practice trials consisted of one example of each condition presented in random order.

Results

The mean response times and error percentages for the 16 conditions are presented in Figure 4. The data were analyzed in two separate repeated measure analyses of variance (ANOVAs). For the response times, the main effects of SOA and intonation were significant, $F(3, 60) = 6.29, p < .001$, $MSE = 16{,}861$, and $F(1, 20) = 14.96, p < .001, MSE = 14{,}018$, respectively. Responses were slower for the 50-ms primes than for the three longer prime durations and slower for targets than for foils.

A significant Intonation × Relatedness interaction, $F(1, 20) = 18.71, p < .001, MSE = 7{,}639$, was modified by a significant three-way interaction, $F(3, 60) = 13.65, p < .001, MSE = 3{,}854$. With the three longer SOAs, the priming effect observed by Bharucha and Stoeckig (1986, 1987) was replicated: Responses were faster to close targets than to distant targets, and there was a tendency in the opposite direction for the foils. With the 50-ms SOA, the effect was reversed, confirming the model's prediction that at short SOAs shared tones should prevail over relatedness.

On the basis of the prediction that with sufficiently short SOAs the priming would be reversed, the interaction of relatedness and intonation was tested for the 50-ms SOA only. This was found to be significant, $t(60) = 2.19, p < .05$, and in the opposite direction of the contrast observed with the longer SOAs. The response times for close and distant pairs were compared with each other for targets and foils separately: Responses were significantly faster for distant targets than for close targets, $t(60) = 1.96, p < .05$, but there was no statistically significant difference for foils, $t(60) = 1.41$.

For the error data, the significant effects were the Relatedness × Intonation interaction, $F(1, 20) = 9.21, p < .01$, $MSE = 177$, and the three-way interaction, $F(3, 60) = 8.18$, $p < .001, MSE = 41$. The direction of the Relatedness × Intonation interaction was reversed with the shortest SOA, but this effect was not significant, $t(60) = 1.04$.

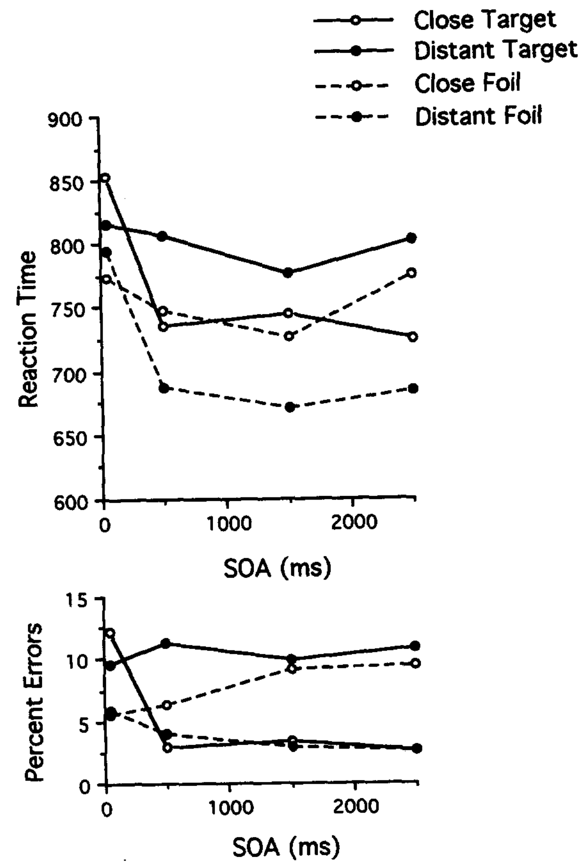

Figure 4. Average reaction times and error percentages for Experiment 1. SOA = stimulus onset asynchrony.

Discussion

Priming was driven by psychoacoustic similarity at 50-ms SOA and by conventional relatedness at SOAs of 500 ms or longer. These results confirm the prediction of the MUSACT model and disconfirm the prediction based on psychoacoustic similarity. We conclude that although both psychoacoustic similarity and conventional relatedness drive priming, the influence of the former is short-lived.

Experiment 2

In this experiment, we tested a contrasting set of close and distant pairs for which the model predicts no crossover. For a C major prime, the close target was D and the distant target was F#. Neither target shares any tones with the prime, so neither chord unit receives direct activation from the tone units in the network. Consequently, both have little activation after the first update cycle (Figure 3). However, they diverge over subsequent cycles because D receives more top-down activation than does F#. As time goes by, D should be primed relative to F# without any reversals.

In an earlier article (Tekman & Bharucha, 1992; Experiment 3), we reported priming as a function of SOA for a close pair with a tone in common (C–G) relative to a distant pair with no tones in common (C–F#). We found priming with an SOA as small as 50 ms. In the present experiment, we chose a close pair with no such advantage over the distant pair. In the first activation cycle, G (the close prime in the earlier article) is strongly activated, whereas D (the close prime in the present experiment) is not activated at all (see Figure 3). We therefore predicted less priming, if any, at 50 ms for D than for G, relative to the same distant target, F#.

The close target in the present experiment is psychoacoustically more similar to the prime than is the distant target, so models based on psychoacoustic similarity predict the same outcome. This is because the close target shares more harmonics with the prime than does the distant target. For example, if the prime chord is C, the second harmonic of tone G (a component of the C major chord) is D, which is the fundamental frequency of tone D in the close target.

Method

Participants. Nineteen undergraduates at Dartmouth College participated in this experiment. All participants reported having normal hearing. The mean number of years they performed with a musical instrument or voice was 5.84, ranging from 0 to 13 years. One participant reported having absolute pitch.

Procedure. The procedure was the same as in Experiment 1 except for the chord pairs used. The roots of the distant pairs were separated by six semitones (an interval of a tritone, or a frequency factor of $2^{6/12}$). The roots of close pairs were separated by two semitones (an interval of a major second, or a frequency factor of $2^{2/12}$) as they were in Experiment 1.

Results

Mean response times and error percentages for each condition are given in Figure 5. In a repeated measures ANOVA for the response times, the main effects of relatedness, $F(1, 18) = 7.76$, $p < .05$, $MSE = 4{,}161$, and SOA, $F(3, 54) = 8.96$, $p < .001$, $MSE = 15{,}751$, were significant. Responses were faster for close pairs than for distant pairs. This replicated the results of Bharucha and Stoeckig (1987) with the same pairs and a 3-s prime duration. The main effect of intonation did not reach significance, $F(1, 18) = 3.84$, $p = .0657$, $MSE = 18{,}909$, but there was a tendency toward faster responses to foils than to targets.

The interaction of SOA and relatedness was not significant, $F(3, 54) = 2.48$, $p = .0706$, $MSE = 2{,}685$, but the effect of relatedness (faster responses for close than for distant) was absent with the shortest SOA. The Intonation × Relatedness interaction was significant, $F(1, 18) = 15.46$, $p < .001$, $MSE = 7{,}806$, replicating the interaction found by Bharucha and Stoeckig (1987), with faster responses for close than for distant targets and slower responses for close than for distant foils.

The three-way interaction was also significant, $F(3, 54) = 6.05$, $p < .001$, $MSE = 3{,}134$, reflecting the different pattern of response times for the shortest SOA compared with the longer SOAs. With the 50-ms SOA, response times for close and distant pairs were not significantly different from each other for targets, $t(54) = 0.79$, and for foils, $t(54) = 0.02$. The difference of these differences was not statistically significant either, $t(54) = 0.57$. Priming was not observed at the 50-ms SOA.

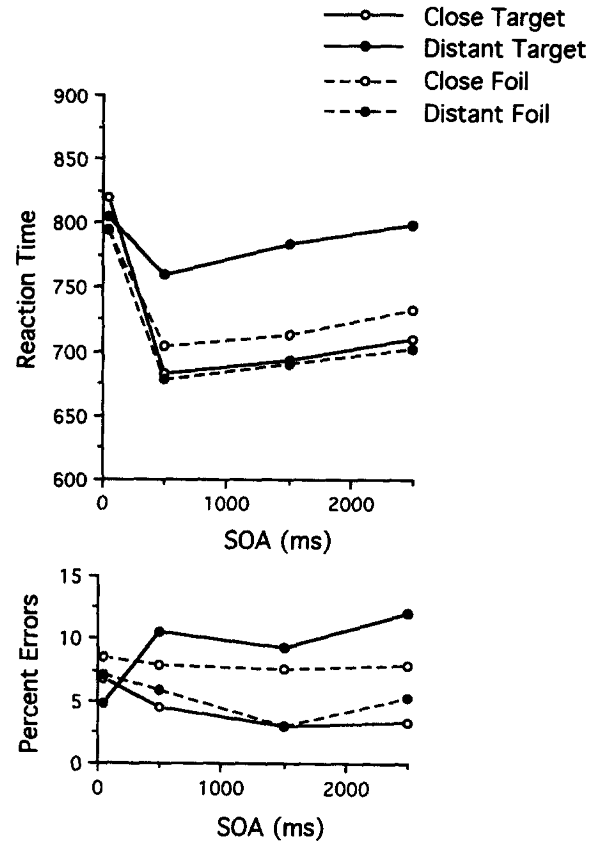

Figure 5. Average reaction times and error percentages for Experiment 2. SOA = stimulus onset asynchrony.

For the errors there were no main effects that approached significance. A significant Intonation × Relatedness interaction, $F(1, 18) = 8.34$, $p < .01$, $MSE = 128$, corroborated the analogous interaction for the response time data. The three-way interaction did not reach significance, $F(3, 54) = 1.83$, $p = .152$, $MSE = 79$.

Discussion

As predicted by the MUSACT model, priming favored the close pairs (C–D) over the distant pairs (C–F#), and there was no crossover as a function of time. Furthermore, no priming occurred at 50 ms, supporting the model's prediction that at the earliest stages of activation there would be less priming, if any, in this experiment (where neither close nor distant pairs shared any tones) compared with Experiment 3 of Tekman and Bharucha (1992), where close pairs shared a tone but distant pairs did not.

General Discussion

The expectations generated by one chord for another are driven by psychoacoustic similarity at very short SOAs and by implicit knowledge thereafter. These results challenge psychoacoustic models that seek to explain musical harmony on the basis of the psychophysics of acoustic spectra alone. They also underscore the importance of dynamic models that can account for the rapid changes in expectation over the course of a single event. The dynamic changes in expectation tested in these experiments were predicted by the MUSACT model, which assumes that pervasive clustering patterns are internalized through perceptual learning. The first-order clustering involves the clustering of pitches to form chords, and the internalization of this structure does not, by itself, add explanatory power over and above psychoacoustic similarity. However, the second-order clustering—of chords in temporal proximity—is based on convention, and the internalization of this structure provides the top-down activation that underlies the eventual dominance of implicit knowledge on chord expectation.

One can speculate that psychoacoustic similarity may have influenced the development of these conventions because the latter reflect the former to a fair extent, and the close correspondence of the two factors has contributed to the long-standing debate about the basis for harmony. However, by considering cases in which psychoacoustic similarity and implicit knowledge of conventional relatedness generate divergent expectations, we have found that conventional relatedness wins out except at very rapid transitions. By 500 ms, conventional relatedness clearly dominates, and most chords (explicit or implied) have even longer temporal windows. These experiments also point to the possibility of making strong, nonintuitive, testable predictions from appropriately constrained neural net models.

It is instructive to consider alternative models that could potentially account for the same data, for example, a two-stage model without any interaction between the stages. The first stage would represent auditory information as spectral information, and psychoacoustic similarity would be the only basis for similarity at this stage. The second stage would incorporate the rules of Western music, thereby representing our implicit knowledge. Assuming that the second stage is engaged more slowly, the model could account for different priming patterns at different SOAs.

In assessing the relative strengths of the two models (MUSACT vs. the alternative two-stage noninteracting model suggested above), the following points are worth making. First, unlike the MUSACT model, the alternative model accounts for the SOA data in a post hoc fashion, whereas MUSACT predicts the data as an unanticipated emergent consequence of its operation. Why, for example, should the second stage of the alternative model be engaged more slowly? In and of itself, however, this does not reduce the plausibility of the alternative model. Second, the connectivity of MUSACT is boot-strapped by a self-organizing system that is exposed to regularities in the musical culture. Although many rule-based systems for Western music have been offered (e.g., Lerdahl & Jackendoff, 1983), not one has been accompanied by a mechanism for learning; the rules are simply assumed to have become represented as a result of an unspecified process. The MUSACT model thus has more explanatory power in the sense that it accounts for how the representation of cultural regularities is acquired.

The experimental designs used in this research did not dissociate prime duration and SOA, raising the question of a possible confound. However, Tekman and Bharucha (1992) did dissociate the two. In that study, SOAs ranged from 50 ms to 2,500 ms. In one experiment, prime duration was identical to SOA. In another experiment, prime duration was held constant at 50 ms, and SOA was varied. No qualitative difference was found in the time course of priming as a function of whether or not SOA was confounded with prime duration: Priming occurred as predicted by the model out to 2,500 ms. If there is any confound between prime duration and SOA in the present article, the confound is biased against our hypothesis. Extending the duration of the stimulus to fill the SOA enhances the sensory effect and diminishes the top-down influence that drives the effect we predict.

Tekman and Bharucha (1992) also showed that the effect of a brief prime can persist over the period of the SOA in spite of masking noise. The results provide, at least, suggestive support for our model's assumption that the flow of activation is phasic, that is, driven by the onset of acoustic energy at any given frequency or by the change in activation at any given node. The salience of tone onsets as compared with the steady-state portions of tones is also supported by auditory physiology and by musical performance practice.

References

Bharucha, J. J. (1987a). MUSACT: A connectionist model of musical harmony. In *Ninth Annual Conference of the Cognitive Science Society* (pp. 508–517). Hillsdale, NJ: Erlbaum.

Bharucha, J. J. (1987b). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception, 5,* 1–30.

Bharucha, J. J. (1991). Pitch, harmony, and neural nets: A psychological perspective. In P. Todd & G. Loy (Eds.), *Connectionism and music* (pp. 84–99) Cambridge, MA: MIT Press.

Bharucha, J. J. (in press). Neural nets, temporal composites, and tonality. In D. Deutsch (Ed.), *The psychology of music* (2nd ed.). Orlando, FL: Academic Press.

Bharucha, J. J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403–410.

Bharucha, J. J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra. *Perception and Psychophysics, 41,* 519–524.

Fukushima, K. (1975). Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics, 20,* 121–136.

Gjerdingen, R. O. (1990). Categorization of musical patterns by self-organizing neuronlike networks. *Music Perception, 8,* 339–370.

Grossberg, S. (1970). Some networks that can learn, remember, and reproduce any number of complicated space–time patterns. *Studies in Applied Mathematics, 49,* 135–166.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding. I: Parallel development and coding of neural feature detectors. *Biological Cybernetics, 23,* 121–134.

Helmholtz, H. L. F. von (1954). *On the sensation of tone.* (A. J. Ellis, Trans.). New York: Dover. (Original work published 1885).

Kohonen, T. (1984). *Self-organization and associative memory.* Berlin: Springer-Verlag.

Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 24–36.

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

Parncutt, R. (1989). *Harmony: A psychoacoustical perspective.* New York: Springer-Verlag.

Rumelhart, D. F., & Zipser, D. (1985). Feature discovery by competitive learning. *Cognitive Science, 9,* 75–112.

Shepard, R. N. (1964). Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America, 36,* 2346–2353.

Tekman, H. G., & Bharucha, J. J. (1992). Time course of chord priming. *Perception & Psychophysics, 51,* 33–39.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America, 55,* 1061–1069.

Terhardt, E. (1979). Calculating virtual pitch. *Hearing Research, 1,* 155–182.

von der Malsberg, C. (1973). Self-organizing of orientation sensitive cells in the striate cortex. *Kybernetic, 14,* 85–100.

Received April 29, 1994
Revision received December 12, 1996
Accepted January 2, 1997 ■