PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Artificial Intelligence and Musical Cognition [and Dicussion]

H. Christopher Longuet-Higgins, B. Webber, W. Cameron, A. Bundy, R. Hudson, L. Hudson, J. Ziman, A. Sloman, M. Sharples and D. Dennett

Phil. Trans. R. Soc. Lond. A 1994 **349**, 103-113
doi: 10.1098/rsta.1994.0116

Artificial intelligence and musical cognition

By H. Christopher Longuet-Higgins

Centre for Research on Perception and Cognition, Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, U.K.

There has been much interest, in recent years, in the possibility of representing our musical faculties in computational terms. A necessary first step is to develop a formally precise theory of musical structure, and to this end, useful analogies may be drawn between music and natural language. Metrical rhythms resemble syntactic structures in being generated by phrase-structure grammars; as for the pitch relations between notes, the tonal intervals of Western music form a mathematical group generated by the octave, the fifth and the third. On this theoretical foundation one can construct AI programs for the transcription, editing and performance of classical keyboard music. A high degree of complexity and precision is required for the faithful representation of a sophisticated pianoforte composition, and to achieve a satisfactory level of performance it is essential to respect the minute variations of loudness and timing by which human performers reveal its hierarchical structure.

1. Introductory remarks

In this paper I hope to convince you that artificial intelligence has made and is still making a useful contribution to the study of musical cognition. By 'musical cognition' I mean the various mental processes that are involved in the perception and performance of music, at least of traditional music . One person's performance is another person's perceptual experience, so the two topics are unavoidably linked.

I shall begin at the perceptual end, by describing some of the problems that arise in the quite ordinary task of transcribing heard music into standard musical notation. Attempts to apply artificial intelligence techniques to this apparently simple skill have led to new insights into the tonal and rhythmic relationships that hold a piece of music together. The second part of the talk will be concerned with the use of computer synthesis for testing ideas about musical performance. I will suggest that performance synthesis enables one to gain a deeper insight into the subtle variations in tempo and dynamic that are essential to the expressive performance of classical piano music.

2. Transcription

A sure sign of musical competence is the ability to transcribe into stave notation a tune one hears played on the piano. This musically unromantic skill is of considerable interest because stave notation conveys much more, and much less, information than just the pitches, loudnesses and onset times of the notes.

Phil. Trans. R. Soc. Lond. A (1994) **349**, 103–113

Printed in Great Britain

TEX Paper

Instead, it reveals the rhythmic and tonal structure of a piece, just as the text of a poem reveals the grouping of the words into lines and stanzas, something that a tape recording conspicuously fails to do.

Unlike composition, where judgments of quality are notoriously subjective, transcription is a task where assessment is all too easy. Figure 1*a* is a transcription of a well-known tune. I suspect you may have some difficulty in recognizing it, but will have less trouble in identifying the tune in figure 1*b*. In fact the two transcriptions call for exactly the same notes on the keyboard, but figure 1*a* totally misrepresents both the rhythmic and the tonal relations between the notes.

Figure 1. Two ways of transcribing the same melody into stave notation, one wrong and the other right. Transcription (*a*) completely disguises the melody by grossly misrepresenting its rhythm and tonality.

This 'right-or-wrong' aspect of music transcription recommends it as a particularly suitable exercise in the computer modelling of musical perception. An obvious comparison is with automatic speech recognition, and one would expect similar problems to arise, such as those due to irregularities of timing. But even when a piece of music is played in strict time, and the notes are identified to the transcriber by their keyboard positions, there is still a quite fundamental problem to be faced at the outset. Every imaginable sequence of notes could be written down in infinitely many different ways without actually misrepresenting their relative lengths or pitches. Consider the extremely simple example of a clock striking 6 (see figure 2). Its chimes could be heard as 2 triplets, or as 3 pairs, or even as occurring halfway between beats. There is no limit to such possibilities, and no obvious way of choosing between them; the enterprise seems doomed to failure at the outset. Undaunted by this prospect, Mark Steedman and I attempted, in the early 1970s, to find ways of putting the barlines into Bach fugue subjects and determining their keys (Longuet-Higgins & Steedman 1971). We found that Her-

Figure 2. Six chimes of a clock, heard in three different ways: (*a*) as two triplets; (*b*) as three pairs and (*c*) as a syncopated sequence.

bert Simon had attempted something similar a few years earlier (Simon 1968), but a principled solution of such problems had to await the importation, into music theory, of some seminal ideas from theoretical linguistics.

3. Generative grammar

In the sense first proposed by Noam Chomsky, the set of all possible melodies in a particular musical style constitutes a language, and to describe a language in finite terms we need a generative grammar (Chomsky 1956, 1965). Various people, including Sundberg & Lindblom (1976), Steedman (1982), Lerdahl & Jackendoff (1983) and Johnson-Laird (1988), have propounded music grammars of various sorts during the last 20 years or so, but none of those authors are to be held responsible for the grammatical ideas I shall now describe (Longuet-Higgins 1976, 1987; Longuet-Higgins & Lee 1982, 1983; Longuet-Higgins & Lisle 1989). The ideas fall under two headings: rhythm and tonality. First, rhythm.

(*a*) *Rhythm*

Informally, the rhythm of a piece of music is the way the notes are grouped together in the time dimension. Take, for example, this little cliché (play figure 3). In stave notation, it looks like this (figure 3). In the language of grammatical

Figure 3. A well-known musical cliché, generated by the metre [2 2 2].

theory, the rhythm of the cliché is generated by the following set of 'realization rules':

(1/2)-unit	→	(1/2)-note	or	(1/2)-rest	or	2 × (1/4)-units
(1/4)-unit	→	(1/4)-note	or	(1/4)-rest	or	2 × (1/8)-units
(1/8)-unit	→	(1/8)-note	or	(1/8)-rest	or	2 × (1/16)-units
(1/16)-unit	→	(1/16)-note	or	(1/16)-rest .		

The symbols on the left of each rule stand for metrical units of various kinds; those on the right include the notes and rests that actually appear on the stave. The first rule says that a (1/2)-unit, such as either of the bars, may be realized as a (1/2)-note, or as a (1/2)-rest, or as two (1/4)-units; the other rules say similar things about the (1/4)-unit, the (1/8)-unit and the (1/16)-unit. Taken together, the first three rules constitute the metre, symbolized by the time signature (2/4) at the beginning of the stave. A more transparent way of specifying the metre would be to write the time signature as [2 2 2]. We can think of this metre as a mini-grammar that specializes in generating binary rhythms.

The rhythm of the cliché is nicely captured by stave notation: the barlines divide it into two (1/2)-units, and the beams joining the notes in the first bar show how they are grouped together, each grouping having arisen from a 'birth of twins' in the generative process. But many other metres are used in traditional music, and each has its own grammatical rules. The metre [2 3 2], for instance, comprises the following rules:

(6/8)-unit	→	(6/8)-note	or	(6/8)-rest	or	2 × (3/8)-units
(3/8)-unit	→	(3/8)-note	or	(3/8)-rest	or	3 × (1/8)-units
(1/8)-unit	→	(1/8)-note	or	(1/8)-rest	or	2 × (1/16)-units.

This metre generates the quite different rhythm of figure 4: different because the (1/8)-units are grouped into threes rather than twos. But notwithstanding this

Figure 4. Another musical fragment, generated by the metre [2 3 2]. The actual note values are the same as in figure 3.

difference in metre, the note lengths in figure 4 are just the same as those in figure 3. This illustrates the important point that a rhythm is not just a sequence of note values, but a logical tree (or a row of trees) representing the temporal relations between successive notes or groups of notes.

In perceiving the rhythm of a melody the listener who has a correct idea of the metre, and of where the barlines come, will be at an obvious advantage over the listener who does not: just like the audience at a symphony concert, who can watch the conductor's movements as they listen to the music. Once the metre is known, and the barlines correctly placed, each note can immediately be assigned a place in the rhythmic hierarchy (though an exception must be made for trills and other ornaments).

The problem of how a listener might infer the metre from the first few notes of a melody still intrigues music psychologists: Peter Desain, Henkjan Honing and I have recently incorporated a simple-minded algorithm for rhythmic perception

Figure 5. Four bars from the last movement of Sibelius's Violin Concerto. Bars 1 and 2 have been notated in 3/4 metre, which implies a syncopation in the first bar but not in the second. Bars 3 and 4 (with exactly the same note values as 1 and 2) have been notated in 6/8; in this metre there is a syncopation in the third bar but not in the fourth.

(after Longuet-Higgins & Lee 1982) into a 'foot-tapping' device that will be put on show later this year (Desain *et al.* 1994).

Though the bar length and the metre are immensely helpful in the rhythmic 'parsing' of a passage, internal evidence may well prompt the listener to revise his/her opinion about the metre or the placing of the barlines. Persistent syncopation, for example, is frequently a sign that the metrical hypothesis needs revision, but the very concept of syncopation was quite vague until it was formalized within the theory of rhythm. Let me give just one example of how the new concept of syncopation illuminates our understanding of a famous piece of music.

In the last movement of Sibelius's Violin Concerto there is a passage that seems to alternate between a 3/4 and a 6/8 metre. On a 3/4 interpretation the rhythms of two successive bars would be as shown in the upper line of figure 5; on a 6/8 interpretation they would be as shown in the lower line. The formal theory of rhythm defines the weight of a note (or rest) as the value of the highest metrical unit that it initiates. On this definition the tied note in the upper line has weight 1/4; its sounded predecessor has weight only 1/8. In the lower line the tied note has weight 3/8; its sounded predecessor only 1/8. In both cases we have a heavier tied note immediately following a lighter sounded note, and a syncopation is formally defined as a sounded–unsounded pair satisfying precisely this condition. So in the 3/4 interpretation the first bar is syncopated, but not the second; in the 6/8 interpretation it is the other way round. This, surely, is why the listener is tilted back and forth between the two metres; a truly exhilarating experience.

Another musical idea that falls straight out of the concept of rhythmic weight is that of metrical accentuation. In the 3/4 waltz rhythm 'dum-diddle-diddle, dum-diddle-diddle' the 'dums' have weight 3/4 and the two syllables within a 'diddle' have weights 1/4 and 1/8. Metrical accentuation stresses the three kinds of syllable according to these relative weights.

(*b*) *Tonality*

Before abandoning perception for performance, a brief word about tonality. If I play the first pair of chords in figure 6 you will hear the first chord as spanning an augmented 4th (play: G, A, B, C sharp); but if I play the second pair of chords, you will hear the first chord quite differently – as spanning a diminished 5th (play: G, A flat, B flat, C, D flat) – even though the two notes concerned (play: G, C sharp) are just the same as before. By supplying a different tonal context the chord has been made to 'flip', rather like the Necker cube so popular in the psychology of vision. By notating the first chord differently in the two contexts

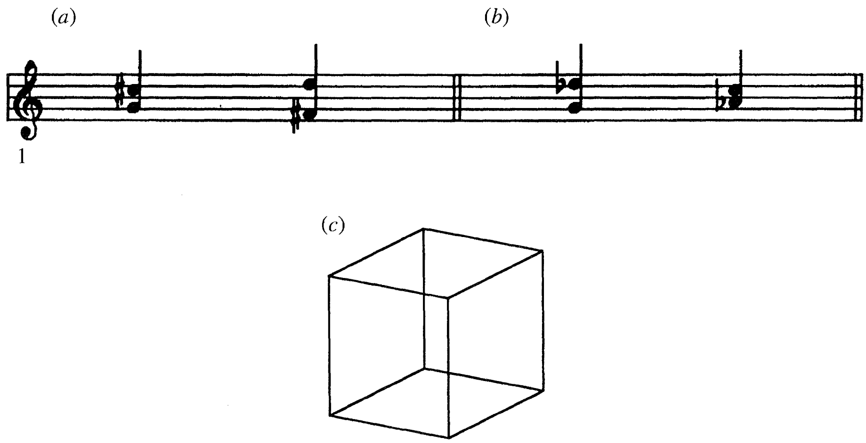

Figure 6. Two chord sequences as they would ordinarily be transcribed into stave notation. On the keyboard the first chord of (*a*) has the same placing as the first chord of (*b*). It used to be called the *diabolus in musica*, and its perceptual ambiguity is analogous to that of the Necker cube (*c*).

the musician is giving overt expression to a perceptual rather than an acoustic distinction. Any transcription program with a claim to psychological reality must obviously do the same. It would not do for such a program to ride roughshod over stave notation, rather as if a spelling reformer were to demand that the words 'hear' and 'here' be spelt in the same way.

To handle the tonal ambiguities that arise in music transcription we need two distinct resources. One is a generative grammar for tonal intervals; the other is a set of heuristics for preferring one tonal interpretation to another in a given context. Tonal intervals are generated, we now realize, not by *division* of the octave into 12 or some other number of equal parts, but by *addition* or *subtraction* of more primitive intervals including the octave, the perfect fifth and the major third. Just as the ambiguity of the Necker cube is associated with a two-dimensional drawing of a three-dimensional object, so tonal ambiguity may be seen as arising from the attempt to project the three-dimensional space of tonal intervals on to the one-dimensional space of the keyboard. But as in rhythm perception, the resulting ambiguities can be resolved by the tonal context, familiar to a musician as the key of the piece. Once a key has been suggested by the first few notes, the following notes virtually fall into place, though, from time to time, the resulting pattern of relations may suggest that a change of key has occurred.

Summing up this bird's-eye view of the transcription problem, we see that its solution depends essentially upon locating the barlines, discovering the metre and the key and keeping the available evidence under constant review. The theoretical

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

framework within which such problems must be formulated was, I think, one of the earliest rewards of applying artificial intelligence to music psychology.

4. Performance

It is now time to turn to my second topic: the contribution of artificial intelligence to our understanding of musical performance, and of piano playing in particular. (I can only hope that musicians will indulge my banal observations on this complex subject!) The human pianist plays from a musical score, which conveys two rather different sorts of information: structural and expressive. The structure comprises all the rhythmic and tonal relations between the notes, and these relations are, as we have seen, essentially discrete. (As a result, classical compositions are potentially immortal, like DNA sequences – an advantage they share with poems but not, sadly, with the murals in the Sistine Chapel!) The expressive information in the score relates, quite indirectly, to the tempo and the dynamics, and how these should vary from moment to moment. I have heard it said that music may consist of notes, but the music is not in the notes, it is in the gaps between them. To demonstrate the truth of this epigram let me play you a technically faultless performance of the first section of Chopin's *Fantaisie-Impromptu* in C sharp minor. In this performance there are no gaps between the notes. Every bar lasts exactly 160 centiseconds, every note has exactly the same loudness and every metrical unit is divided into subunits of exactly equal duration. (Play tape A.)

While we recover from that rather distressing experience, let me say a word about the information technology involved in producing even such a drab performance. One needs an electronic keyboard and a computer that can either take information from the keyboard or issue instructions to it in real time. There are three programs stored in the computer: a Parser, an Editor and a Performer. The user plays the music on the keyboard, slowly and painfully, and the Parser converts it into a data structure that makes explicit the rhythmic relations between the notes. The Editor is an interactive program that permits the user to amend or otherwise modify this data structure – let us refer to it as the score – and to augment it, as a human editor might, with indications of tempo, dynamic, articulation and – most important – expressive variations of tempo. Finally, the Performer converts all this information, by rule, into a data stream of note events that are realized in sound by the keyboard. It is here that the heart of the system really lies.

Though essential for aesthetic reasons, expressive gestures present two immediate problems. First, there is the problem of deciding how a *crescendo* or a *rallentando* should be physically realized. But worse, many of the gestures that a composer or performer would regard as appropriate are not marked in the score at all. The latter problem can be circumvented by heavy editing of the score in one's capacity as a literate musician; but the former problem – that of the forms and parameters of musical gestures – can only be determined by careful research, such as that of Bruno Repp (1990) in the United States, Johan Sundberg (1988) in Sweden, and Eric Clarke (1985) and Neil Todd (1985, 1992) in England.

In a detailed study of Schumann's *Traümerei* as played by a number of famous pianists, Repp (1992) has found that they all impose smooth, convex tempo contours on the quaver sequences, speeding up at the beginning, and slowing down at the end of such a sequence. Adopting a rather different approach, Sundberg and his colleagues (Sundberg 1988; Sundberg *et al.* 1991) have plucked a number of plausible rules out of the air, as it were, and used them in the production of synthetic performances. A panel of sophisticated listeners is then invited to judge the relative merits of these performances, and the rules are 'fine-tuned' accordingly. A number of such rules have emerged, mostly about the realization of single melodic lines.

I am inclined to agree with Sundberg that the best, if not the only, way of establishing the musical effectiveness of a particular type of gesture is to judge how its use affects the quality of a synthetic performance. But rather than attempting to implement Sundberg's own rules, my strategy has been to try to follow the fairly plentiful marks of expression that actually appear in authoritative editions of the classical masters. Some, such as the accents and other dynamic markings, turn out to be relatively unproblematic, and some, such as *molto agitato*, seem to happen of their own accord if one observes other, more detailed, indications. But this is emphatically not true of the temporal shaping of phrases: the groups of notes or bars delimited by 'slurs' in the scores of classical pianoforte works. In their pioneer studies of expressive timing, Shaffer, Clarke & Todd found that pianists playing pieces by Satie and by Chopin marked the boundaries of each phrase by a tempo contour of the same general shape as the slur that delineates the phrase in the score (Shaffer *et al.* 1985; Clarke 1985; Todd 1985). The tempo rose at the beginning of the phrase and fell towards the end, and the same was true of shorter phrases inside longer ones.

To convert such a discovery into a computational procedure, one has to find, by trial and error, the parameters that relate the phrase structure with the tempo contour. In a moment I shall play another synthetic performance: this time of the third section of the *Fantaisie-Impromptu*. In this performance the gesture of phrase-final tempo relaxation has been deployed at all levels from the bar itself up to the 12-bar or 16-bar phrase. Apart from two small tempo changes during the passage marked *poco a poco più tranquillo*, the tempo is a continuous function of time, and linear within each half-bar. At the end of each phrase it falls to a value proportional to the length of that phrase, but this fall is anticipated by one of half the size, half a bar earlier. In the closing bars, the tempo (measured in semiquavers per second) is strictly linear in the time from the beginning of the *ritardando* to the end of the last note, at which point it has fallen to zero. As for other expressive gestures, they are kept to a bare minimum to avoid distracting the listener from the phrase structure. There are six dynamic levels, ranging from *pianissimo* to *fortissimo*, and each hand maintains a constant level (apart from accents) through each half-bar. Here, then, is the concluding section (bars 83–138) of Chopin's *Fantaisie-Impromptu* in C sharp minor, realized in this way. (Play tape B.)

5. Concluding remarks

There is no need to draw attention to the great gulf between these music programs and the human faculties to which they might be relevant. Nevertheless I believe that work of this kind has enhanced, and will continue to enhance, our appreciation of the musical mind. It has already inspired the development of a formal theory of rhythm, and justified earlier developments in the theory of tonality. It has shown how rhythmic and tonal relations can be efficiently represented by symbolic structures that make explicit such concepts as key, metre, rhythm, bar, beat, phrase and many others that are essential to the coherent discussion of classical music.

Finally, a brief comment on the issue of symbolic versus sub-symbolic models of mental processes. It would be difficult to exaggerate the computational advantages that have flowed from representing pieces of music by structures that contain exactly the information one finds in a musical score. Whether they are so represented in our own minds is, of course, an open question; but concert pianists would quickly come to grief if their musical memories were corrupted, even very slightly, by the kind of cross-talk that plagues neural networks and other sub-symbolic models of cognition. I have a feeling that cognitive models of the well-tried symbolic variety will be with us for a long time to come.

I thank my colleague Edward Lisle for his invaluable technical assistance.

References

Chomsky, N. A. 1956 *Syntactic structures.* The Hague: Mouton.

Chomsky, N. A. 1965 *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

Clarke, E. F. 1985 Some aspects of rhythm and expression in performances of Erik Satie's 'Gnossienne No. 5'. *Music Perception* **2**, 299–328.

Desain, P., Honing, H. & Longuet-Higgins, H. C. 1994 A musical shoe. In *Proc. Int. Computer Music Conf.* (In the press.)

Gabrielsson, A. 1987 The theme from Mozart's piano sonata K.331. In *Action and perception in rhythm and music* (ed. A. Gabrielsson), publ. 55, pp. 81–104. Stockholm: Royal Swedish Academy of Music.

Johnson-Laird, P. N. 1988 *The computer and the Mind.* Cambridge, Mass.: Harvard University Press.

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

Longuet-Higgins, H. C. 1976 The perception of melodies. *Nature, Lond.* **263**, 646-653.

Longuet-Higgins, H. C. 1987 *Mental processes*, ch. 7–13. Cambridge, Mass.: MIT Press.

Longuet-Higgins, H. C. & Lee, C. S. 1982 The perception of musical rhythms. *Perception* **11**, 115–128.

Longuet-Higgins, H. C. & Lee, C. S. 1983 The rhythmic interpretation of monophonic music. In *Studies of music performance* (ed. J. Sundberg), publ. 39, pp. 7–26. Stockholm: Royal Swedish Academy of Music.

Longuet-Higgins, H. C. & Lee, C. S. 1984 The rhythmic interpretation of monophonic music. *Music Perception* **1**, 424–441.

Longuet-Higgins, H. C. & Lisle E. R. 1989 Modelling musical cognition. *Contemporary Music Review* **3**, 15–28.

Longuet-Higgins, H. C. & Steedman, M. J. 1971 On interpreting Bach. In *Machine intelligence*, vol. 6 (ed. B. Meltzer & D. Michie). Edinburgh University Press.

Repp, B. 1990 Patterns of expressive timing in performances of a Beethoven minuet by 19 famous pianists. *J. acoust. Soc. Am.* **88**, 622–641.

Repp, B. 1992 A constraint on the timing of a melodic gesture: evidence from performance and aesthetic judgment. *Music Perception* **10**, 221–242.

Shaffer, L. H., Clarke, E. F. & Todd, N. P. 1985 Rhythm and metre in piano playing. *Cognition* **20**, 61–77.

Simon, H. A. 1968 Perception du pattern musical par AUDITEUR. *Sciences de l'Art* **5**, 28–34.

Steedman, M. J. 1982 A generative grammar for jazz chord sequences. *Music Perception* **2**, 52–77.

Sundberg, J. 1988 Computer synthesis of musical performance. In *Generative processes in music* (ed. J. Sloboda), pp. 52–69. Oxford: Clarendon Press.

Sundberg, J., Friberg, A. & Fryden, L. 1991 Threshold and preference quantities of rules for music performance. *Music Perception* **9**, 71–82.

Sundberg, J. & Lindblom, B. 1976 Generative theories in language and music descriptions. *Cognition* **4**, 99–122.

Todd, N. P. 1985 A model of expressive timing in tonal music. *Music Perception* **3**, 33–35.

Todd, N. P. 1992 The dynamics of dynamics: a model of musical expression. *J. acoust. Soc. Am.* **91**, 3540–3550.

Discussion

B. Webber (*University of Pennsylvania, U.S.A.*). Using a linguistic analogy, the phrase structures marked on Professor Longuet-Higgins's score seem to correspond to paragraphs. But the phrases in the pieces are often interrupted. So the analogy isn't straightforward.

C. Longuet-Higgins. There is a hierarchy of phrases in a piece. The marked phrases are at a fairly high level. Someone editing a score must decide which phrases it would be most useful to mark. If every phrase were marked, the score would become too cluttered. The rest of the phrasing, the phrases within phrases, is evident to a trained musician and doesn't need to be marked explicitly. In the Chopin *Fantaisie-Impromptu* you just heard, the heuristic concerning cumulative delay proportional to phrase-length is deployed at all levels. So yes, there are interruptions. Whenever a subphrase is completed, the larger phrase is picked up again.

W. Cameron (*University of Glasgow, U.K.*). With poetry and music, a temporal component is crucial: they must be apprehended in real time.

C. Longuet-Higgins. Poetry and music have much in common. Linguists tend to be preoccupied with texts, but we shouldn't forget performance. Verse is meant to be read out loud, declaimed. Nor should we forget that setting words to music can be done well or badly, and there are fairly objective criteria here. Painting is different: it doesn't have a temporal dimension. But cinema does.

A. Bundy (*University of Edinburgh, U.K.*). Professor Longuet-Higgins mentioned temporal changes exploited for marking phrases. Are similar rules used in spoken language?

C. Longuet-Higgins. Yes. Written language has various marks of punctuation which indicate phrase markings, just as the score does, but not all of the phrase markings are explicitly written down. In speech, phrases are marked by temporal changes and other markers. A good speaker is particularly adept at this.

R. Hudson (*University College London, U.K.*). Why is musical sensitivity distributed very unevenly through the population, when compared with the remarkably even spread of linguistic sensitivity?

C. Longuet-Higgins. We're not all equally sensitive to Chinese! It's just a matter of learning.

L. Hudson (*University of Wales, Cardiff, U.K.*). Playing Bach, you use much less *rubato* (local changes of tempo) than when you are playing Chopin or Liszt. Are different rules used for marking the phrases?

C. Longuet-Higgins. You keep a tight rein on tempo changes when playing Bach. In all cases it is the same rule operating, but for different styles of music you set different parameters. By varying the parameters you could even imitate the different playing styles of various performers.

J. Ziman (*Epistemology Group, Aylesbury, U.K.*). Can some people go to a performance and write down a score? Or do different people come out with different answers?

C. Longuet-Higgins. For tunes, people can do it, and they come up with the same answer. But with a Bach fugue, a listener can't capture the complete score from one listening. There's just too much going on, and one can't attend to it all.

A. Sloman (*University of Birmingham, U.K.*). Is the Sibelius piece you mentioned uniquely transcribeable?

C. Longuet-Higgins. Sibelius actually marks two distinct time signatures in the score.

M. Sharples (*University of Sussex, U.K.*). Could Professor Longuet-Higgins's techniques apply to moving images?

C. Longuet-Higgins. People making cartoons often start with the soundtrack, and then synchronize the images so that important visual events coincide with musical events. There's a good example in Disney's *Fantasia*, where bursting mud-bubbles are synchronized with dramatic chords from Stravinsky's *Rite of Spring*.

D. Dennett (*Tufts University, U.S.A.*). In composition competitions where competitors submit scores, not tapes, the judges use only the scores. Just how 'high-fidelity' is their reading of them? Would they notice a small deliberate mistake?

C. Longuet-Higgins. Professional score readers, such as people who mark composition exercises, are remarkably astute, and do notice very subtle details. They might not notice some blemishes in the score, just as you may overlook typographical errors. You substitute the correct reading, based on your knowledge of the context. So do they.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES