Esa Itkonen

On (Sign) Language, Music, and Anti-Modularity

There are two basic arguments against the view that language is a mental module. First, language is not encapsulated vis-à-vis the extralinguistic reality, as shown by *iconicity*, or the structural similarity between the two realms. Second, language is not encapsulated vis-à-vis other mental domains, as shown by the fact that the same structures and processes obtain e.g. in language, spatial perception, logic, and music. The most important of such pervasive mental capacities is *analogy* (taken either in the static or in the dynamic sense). A significant generalization is achieved once it is realized that iconicity itself is just a particular instance of analogical thinking: 'structural similarity', which was stated above to obtain between language and the extralinguistic reality, is the *definition* not just of iconicity, but also of analogy in general (for discussion, see Itkonen & Haukioja 1997).

The case for analogy (and, thus, for anti-modularity) cannot be fully appreciated as long as the existence of iconicity is ignored. The current 'imagery debate' illustrates this point quite well. One side, led by Pylyshyn, argues that thought is digital or language-like, whereas the other side, led by Kosslyn, argues that thought is picture-like (cf. Tye 1993). Let L, T, and R stand for 'language', 'thought', and 'reality' (i.e. 'reality-as-perceived'); let A → B mean 'A represents B', and let A ⇒ B mean 'A produces the structure of B' (or 'B reproduces the structure of A'). Then the two positions of the debate may be presented as follows:

$$\text{Pylyshyn:} \quad L \to T \to R$$
$$\Rightarrow$$

Kosslyn: L → T → R
                      ⇐

A few comments are now in order. First, the tripartite relation L → T → R, on which both sides agree, follows exactly the Aristotelian line of thinking (cf. Itkonen 1991: 175-176). Second, Pylyshyn's position is more ambiguous than it seems, because he is committed to the view that thought is a *digital picture* of language. Third, the possibility of L ⇐ R has been overlooked in this debate. Iconicity shows, however, that language is a picture of reality. Therefore the whole opposition 'language-like vs. picture-like' collapses, at least to some extent. (That is, language is an abstract or structural picture, not a concrete picture on a par with ordinary mental images.) The foregoing presupposes that if there is a non-digital relation between reality-as-perceived and language, then the 'intervening' level of thought cannot possibly be digital.

My concern here is with the modularity thesis. In my previous publications I have discussed the nature of iconicity (cf. Itkonen 1994), and I have also explored the similarity between linguistics and formal logic (cf. Itkonen 1978: chap. 10). In this paper I intend to have a look at the relation between language and music.

Jackendoff (1987: chap. 11) offers a convenient starting point. He distinguishes between four distinct levels of musical structure. For instance, the opening line of Mozart, K. 550 has 20 notes, which constitute a 'grouping structure' in the following way (cf. Fig. 1): notes 1-3 (= A), 4-6 (= B), 7-10 (= C), 11-13 (= D), 14-16 (= E), and 17-20 (= F) are constituents of the lowest level; constituents of the next level are formed by A and B (= G) as well as by D and E (= H); constituents of the next level are formed by G and C (= I) as well as by H and F (= J); the one-line-length unit itself is constituted by I and J and will in turn be a constituent in higher-level units. The constituents are held together by the standard gestalt criteria, i.e. similarity and/or proximity between units, and are — inversely — separated by the opposite criteria, i.e. dissimilarity and/or distance.
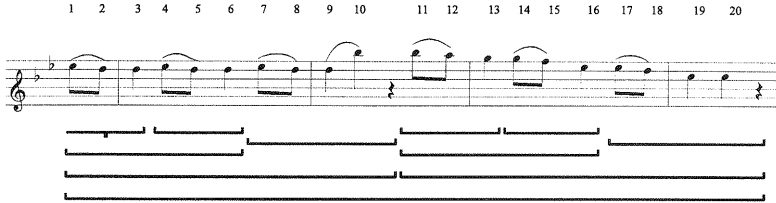
Fig. 1. Grouping structure in the opening of Mozart's G minor symphony, K. 550 (from Jackendoff 1987, 219).

Next, there is a 'metrical structure' consisting of strong and weak beats. The 20-note line mentioned above contains 32 beats, as follows: 21412131512131212141213121512131. The 'strength' of a beat is constituted by a heavy stress and/or by a relatively long duration. There is a tendency to put the strong beat at the beginning of the units of the grouping structure. Just like the grouping structure, the metrical structure too may be represented with the aid of a tree diagram.

The level of 'time-span reduction' introduces the distinction between a theme and its variation (or elaboration). The former represents what is important; it is the 'skeleton' enriched by the latter. The enrichment may be either sequential (= melody) or simultaneous (= harmony). At this level too, structures may be depicted with the aid of tree diagrams, with the qualification that the 'theme vs. variation' distinction becomes visible only by comparing *several* trees simultaneously.

Finally, the level of 'prolongational reduction' introduces the distinction between the building-up of tension and its relaxation. It represents the 'movement' within a piece of music. Correspondingly, while the elaboration of units at the previous level is symmetrical, it is asymmetrical at this level.

It is quite easy to see the analogy (sic!) between the musical structures and the linguistic structures. The grouping structure

strongly resembles the constituent structure of a sentence. The fact that the same type of structure seems to be "involved in any sort of temporal pattern perception [and production!]" (Jackendoff 1987: 221), does nothing to diminish the importance of this overall similarity. Rather, it constitutes a general argument for anti-modularity. The metrical structure is practically the same in language and in music; for instance, Jackendoff (1987: 79) analyzes the metrical structure of the expression *American history teacher* as 121141131, in much the same way he analyzes the opening line of Mozart, K. 550 (cf. above). At the level of time-span reduction, the distinction between theme and variation closely parallels that between obligatory and obligatory-*cum*-optional constituents (cf. *The boy ate an apple* vs. *The little boy hastily ate a red apple*); and both in language and in music, the obligatory constituents are the 'heads' of their respective domains. It is curious that Jackendoff does not point this out explicitly, especially since Sapir (1921: 36), for instance, applied the same term, i.e. 'reduction analysis', to the method of finding out the obligatory constituents of a sentence. (It may be added that this method was already practised by Apollonios Dyskolos, the first syntactician of the Western tradition; cf. Itkonen 1991: 202-203). Finally, the distinction between the levels of time-span reduction and prolongation reduction parallels that between the levels of sentence and text (or discourse). Again, Jackendoff fails to mention this obvious fact.

While Jackendoff (1987: chap. 11) is content to analyze the structural aspect of music, he turns to the question of how music is processed (i.e. heard and understood) in chapter 7 (= 'Musical parsing and musical affect') of his 1992 book. He notes explicitly (p. 125) that he will proceed "by analogy with evidence from the processing of language", a problem that he had adressed in chapter 6 of his 1987 book. His solution is, briefly, that both language and music are understood by constructing parallel interpretations, from among which one will ultimately prevail. This is certainly plausible.

However, Jackendoff's overall argument contains the following oddity. On the one hand, he uses continuously such words as 'analogy', 'similarity', and 'parallelism'; and he clearly assumes that the plausibility of the hypothesis that language is processed in a parallel fashion *supports* the analogous hypothesis about how music is processed. On the other hand, as a "deeply committed Chomskian" (1992: 53), he accepts Chomsky's view that language and (e.g.) music are separate mental modules. But this is inconsistent: if language and music are independent of each other, then any similarity between the two is due to *chance*; and results achieved in one domain cannot possibly support hypotheses about the other domain.

In other words, Jackendoff takes the both structural and processual analogy between language and music in a purely *heuristic* sense: language just happens to illuminate music, and vice versa. It must be asked, however, what is the basis of this pervasive analogy. It is just too implausible to think that it is a result of pure chance, or that it rests on some sort of 'pre-established harmony'. The only reasonable answer is that language and music emanate from a common source, or instantiate a more general capacity; that is, the analogy must be taken in an *ontological* sense. Because of his a priori commitment to the Chomskian modularity, Jackendoff is prevented from drawing this obvious conclusion. (To be sure, he assumes the existence of 'general properties of the computational mind', but the implications that this assumption has for the modularity thesis are never spelled out.)

All in all, the evidence presented by Jackendoff speaks strongly *for* the similarity of language and music (and thus against any modular interpretation of the relation between the two). Therefore, if we wish to find some evidence *against* the similarity of language and music, we must look elsewhere.

According to Thomas (1995: 12-13), the 20th-century musical theory has in general disregarded the possibility that music might have any referential or expressive functions, and has instead posited the existence of 'pure' or 'absolute' music. If we are

interested in elucidating the relationship between music and language, it is not very meaningful to consider a position which simply ignores the possibility of any such relationship. A more fruitful position in this respect is represented e.g. by Davies (1994). On the one hand, he rejects the notion of 'absolute music', because he assumes that music has at least some sort of relation to such a 'substantive' notion as *feeling*. On the other hand, however, he argues at great length against the view that music could be interpreted as (analogous to) a *symbol* of any kind, be it linguistic, pictorial, metaphoric, or 'representational'. According to him, there are first of all several reasons why music cannot be language-like: it lacks the distinction between illocutionary force and propositional content, i.e. it neither asserts (truly or falsely) nor asks nor commands; it lacks (counterparts to) logical connectives; it has no (counterpart to) metalanguage; etc. Second, music is not picture-like, because — quite obviously — no concrete things or events can be 'recovered' from a musical performance with anything approaching intersubjective agreement. Third, music is not metaphoric, because while metaphors can in general be paraphrased in non-metaphorical terms, this cannot be done to a statement like 'This piece of music is sad'.

Up to this point, it is easy to agree with Davies. However, this has hardly any bearing on the modularity issue. It would be quite unrealistic to demand a low-level point-by-point similarity between language and music. And Davies's general approach is so 'philosophical' that he never comes to grips with the systematic structural-*cum*-processual similarities between language and music that have been pointed out above. Thus, the argument for anti-modularity, or the 'common source' of language and music, remains intact.

Nevertheless, it is meaningful to ask whether Davies (1994) is capable of throwing any additional light on the issue. The opportunity to do so arises when he comes (pp. 123-134) to examine Langer's (1942) philosophy of music. According to Langer, language and music are similar in being symbolizations, but they are different insofar as language is a 'discursive' symbol

while music is a 'representational' symbol. A representational symbol is supposed to symbolize the form common to all feelings (or, alternatively, the form common to the various instances of a particular feeling). It is based on the *iconicity* of music/art and feeling insofar as a sameness is experienced between the forms of works of art and the forms of feelings.

Davies rejects Langer's theory, on two accounts. First, the concept of representational symbol presupposes the view of the iconicity of language (as represented e.g. by the 'picture theory' of the early Wittgenstein); but — according to Davies — this view has subsequently been abandoned by everybody (including Wittgenstein himself). Second, because the nature of the representational symbol cannot be discursively discussed, it is simply obscure (or incoherent).

Davies's own view is as follows. Music neither symbolizes nor expresses feelings. Rather, in musical contexts we act just like we do in those "nonmusical contexts in which we attribute emotion characteristics to the appearances of people, or nonhuman animals [cf. the 'sad look' of Basset hounds], or inanimate objects [cf. the shape of the weeping willows which resembles the shape of people who are downcast and burdened with sadness]" (p. 228). This view is indebted to Wittgenstein's later philosophy which denies any strict separation between 'inner' feelings and 'outer' behavior.

Davies's view calls for some comments. First of all, as is evident from our discussion of the 'imagery debate', it is absolutely wrong to claim that iconicity of language is an outdated doctrine. More specifically, it is wrong to claim that the 'use theory of language' of the later Wittgenstein contradicts his earlier 'picture theory'. Rather, the former simply contains the latter (see e.g. Kenny 1973: chap. 12).

Secondly, Davies is quite right to point out the simple-mindedness of the notion that, on one side, there are the feelings and, on the other, there is that by which they are expressed (be it language, music, or ordinary behavior). Wittgenstein's private-language argument has demonstrated the impossibility of inner mental states that have no systematic public criteria (cf. Itkonen

1978: 4.2). In the current metatheory of psychology, including psycholinguistics, this simple-minded separation of mind and behavior is advocated by the 'Neo-Cartesian' school, represented most prominently by Chomsky's generative linguistics (cf. Itkonen 1983: 5.1). It is not without interest to note that Jackendoff, a "deeply committed Chomskian" (cf. above), must be counted among the Neo-Cartesians.

From the fact that the levels of ('inner') feeling/thought and ('outer') behavior are conceptually interdependent, it does not follow, however, that they are identical. It is a matter of conceptual necessity that, in order to exist, a feeling or thought requires *some* public criterion *most* of the time; but the precise nature of this criterion is unpredictable, and sometimes it may be lacking. Therefore it is still admissible, *grosso modo*, to speak of feelings/thoughts and their expressions as if they were two distinct levels. In this way it seems possible to retain Langer's (1942) analysis, while doing justice to Davies's (1994) misgivings: sadness is attributed to certain characteristics of music just like it is attributed to certain characteristics of the human face; but in both cases, these characteristics do express the corresponding feeling, even if in a 'conceptually dependent' way. (By contrast, the 'sadness' of dogs and willows is metaphorical.)

Thus, the analogy between language and music seems to hold up rather well. But now an important qualification has to be made. When the term 'language' is used in this context, it is always *oral* language which is meant. This seems natural enough because music too is concerned with *sound*. It is generally acknowledged today, however, that the sign languages of the deaf are on a par with the oral languages. This means that the traditional notion of language has to be revised accordingly, i.e. it has to be made abstract enough to subsume languages in the oral (= auditory) mode and in the signed (= visual) mode as particular instances. And the next step will be to subsume the tactile languages of the deaf-and-blind in the same way.

This revision has the interesting consequence that from now on, while discussing music, we have to regard its auditory aspect as

merely incidental. (To be sure, it could be argued that this change of perspective is not all that dramatic because both speech and music can be written and thus already exist in the visual mood.) The following structural levels remain intact: a) grouping structure as the general hierarchical patterning of (symbolic) actions; b) the distinction between a 'core structure' consisting of obligatory elements and its elaborations that add optional elements; c) the distinction between 'sentence' and 'text' (or 'event' and 'story'). By contrast, the 'metrical structure' has to be reinterpreted so as to eliminate any reference to sound. What remains, is the distinction between emphatic and less emphatic; and it may be conveyed by a corresponding auditory distinction (cf. above) or visual distinction (= large and/or quick motion vs. its opposite) or tactile distinction (= more vs. less pressure).

Thus, just as there is an analogy between music and spoken language, there is also an analogy between music and sign language. It may be claimed, in addition, that sign language is not just analogous to music, but also has a 'music' of its own. (Hence there are in fact two analogies instead of just one.) Of course, the notion of 'music without sound' may seem preposterous at first. However, it can be made more comprehensible by pointing out that an expression like 'song in a language without sound' actually occurs in the title of a 1976 paper by Klima and Bellugi. They analyze the changes that are involved in the *artistic* use of American Sign Language, as compared with its everyday use. First, they observe three types of change that might be characterized as 'poetic': choice of vocabulary; increased tendency to use both hands; tendency to build uninterrupted transitions between particular signs, thus creating an impression of a 'flow of movement'. Even more interestingly, however, they also observe a change that they characterize as 'musical': the movements of the two hands are much enlarged, and they are placed in such a way as not to intersect. According to Klima and Bellugi (1976), this last-mentioned pattern is imposed upon signing just like, in a song, the melody (= one type of sound structure) is imposed upon the words (= another type of sound strucure). (To be sure, it could be argued

that there is something musical already in the attempt to emphasize the 'flow of movement' aspect of signing.) — It may also be added that there is a striking similarity between the activities of conducting a symphony orchestra and conducting a chorus of signers.

One complication still has to be mentioned. In singing, it is always possible to eliminate the words and leave the music, which — presumably — exhibits the four types of structure mentioned above. It is less clear whether there actually exists a comparable practice of 'purely musical' signing (or 'humming') among the users of sign language. Some informal reports that I have received do, however, confirm this possibility.

Up to now, I have been trying to show that there are plausible similarities between music and spoken language, on the one hand, and between music and signed language, on the other; and such similarities can be interpreted as pointing to a common source. Language has referential meaning whereas music lacks it. My argument would be strengthened if it could be shown that between music and (typical) language there is such a missing link as 'language without referential meaning'. Now, it is the central claim of Staal (1989) to have discovered precisely this type of phenomenon in the mantras of the Vedic age. They were uttered or performed according to strictly defined rules, but they just had no meaning, at least not in the referential sense. More obviously, the same phenomenon occurs in singing, because it is self-evident that singing often involves words (or 'words') with no or very little meaning.

Comparing language with music is apt to de-emphasize its referential aspect and to emphasize its *rhythmic* aspect. Moreover, the combination of rhythm and *iconicity*, as it occurs in Langer's (1942) theory of music, provides a new perspective on the iconicity of language. In rhythm, there is an inseparable connection between doing and understanding (or responding); one cannot be without the other. A notion like 'rhythmic iconicity' presupposes that there is rhythm not only in the symbol but also in that which is symbolized. Now, this is *not* how iconicity is generally understood

in linguistic theory. In it, a sentence is taken to be a (structural) picture of a state of affairs; but the picturing (which inevitably has its own rhythm) is thought to be performed always alike, independently of whether that which is pictured is a man hitting a dog or the sun setting behind a mountain. But now the notion of rhythmic iconicity suggests that the primary case might be the one where there is doing (and understanding) both on the 'picturing' side and on the 'pictured side': one can picture only what one has done or can do oneself.

The foregoing could be taken to stress the metaphorical aspect of language (and thought): human action is primary, and everything else is understood on the model provided by it. However, this interpretation might be a little hasty. (Remember that the cave men also painted animals in at least apparent isolation either from men as such or from men hunting animals.) What this interpretation really points to, is something more subtle: the reality as a whole, i.e. both its human and its non-human aspect, has been equally conceptualized by man. Therefore we possess an intimate knowledge of this conceptualization. It is this truth that the doctrine of 'response-dependence' has recently rediscovered (cf. Itkonen 1978: 42-43, 1997: 59-60). The rhythm that language pictures is not that of actions (like a man hitting a dog), understood as antithecal to physical events, but that of conceptualizations in general, be it conceptualizations of actions (like a man hitting a dog) or of physical events (like the sun setting behind a mountain). — It is vital to understand the following distinction: the sun has not been made by man, but the concept 'sun' has.

In what precedes, I have argued for the view that language and music derive from a common source. This statement is meant to be taken in the *synchronic* sense, i.e. as being about the human mind as it exists now. It is tempting, although in no way mandatory, to take the statement also in the *diachronic* sense, i.e. as being about the (evolutionary) origin of language and music. In the present context I shall resist this temptation only in part and thus add a few words on the latter topic as well.

Staal (1989) was mentioned above as providing evidence for the synchronic commonality of language and music. However, he personally regards his analysis of the 'language without meaning', exemplified by the Vedic mantras, as a contribution to clarifying the *origin* of language. As he sees it, this type of non-referential, and hence 'musical', use of structured sound was the precursor of referential language: meaning was secondarily imposed upon sound that was already there. Of course, this sound too had to have some sort of function, but this was to express or sustain hard-to-define communal feelings. Wallin (1991) reaches a similar conclusion from a rather different, i.e. biological and cross-species, point of view. According to him, language was, and still is, prefigured by the calls emitted by non-human vertebrates. He assumes a slow evolutionary process during which sound patterns acquire more and more differentiated meanings. More recently, similar views have been expressed e.g. by Bruce Richman, Dean Falk, and Björn Merker. They all assume that, to put it very simply, language emerged from singing. This type of vocalization was performed in a group (and still is among certain types of apes).

The collective nature of this pre-language explains quite naturally how meanings, once they emerged, could be learned by all members of the community simultaneously. In this type of group performance, exemplified equally by song and dance, every participant has to *synchronize* his actions with those of every other participant. Consequently, behavior of this type literally *embodies* the notion of common knowledge (= 'everybody knows-1 X and knows-2 that everybody knows-1 X'). The theoretical importance of song and dance becomes evident when we recall that *language exists only as an object of common knowledge* (cf. Itkonen 1978: 122-131, 1997: 54-57; Clark 1996: 93-95). Keeping in mind the analogy with song and dance, we are able to grasp better than before the sense in which, in a typical verbal exchange, the speaker and the hearer have to synchronize their respective actions in relation to what has been and is being said. (The term 'synchronize' is more adequate than the fashionable term 'negotiate'.)

It is quite interesting to note that these modern views were anticipated in great detail by Jespersen (1922: 392-442). In his speculations about the origin of language, he arrived at the view that referential language must have been preceded by singing, which in its turn was functional in fulfilling the need for sex (or love), on the one hand, and the need for coordinating collective work, on the other.

Going still farther back in history, we come across other similar anticipations. In antiquity, it was generally thought that people first discovered music in the nature that was *external* to them, and then invented ways to imitate it. It was assumed, for instance, that "the imitation of bird-calls led to the first songs", or that instrumental music originated as an imitation of "the sounds the wind produced in reeds growing on the banks of the Nile" (Thomas 1995: 50). There exists a more elaborate story about how the notions of pitch and harmony were discovered: "Pythagoras, according to tradition, is supposed to have discovered that musical pitch depends on the ratio between the length of vibrating chords — the starting point of mathematical physics — by passing in front of the local blacksmith on his native island of Samos, and noticing that rods of iron of different lengths gave different sounds under the blacksmith's hammer" (Koestler 1967: 111). This view, that harmony is a physical phenomenon produced by a resonant body (*corps sonore*), was central to the thinking of Jean-Philippe Rameau, the most celebrated musical theorist of the 18th century.

By contrast, the philosophers of the French Enlightenment generally placed the origin of music in the human voice, thus making it *internal* to human beings. Music and language were thought to have emerged from a stage at which the distinction between the two did not yet exist. This view was held both by Condillac and by Rousseau, although they articulated it rather differently. In his *Essai sur l'origine des connaissances humaines,* Condillac envisages the evolution of language as a three-stage process. At first, there was a close cooperation between voice and gesture: 'natural cries' (*cris naturels*) expressed the fact that one was deprived of some object that one wanted to have, and gestures

indicated which object it was. Next, the vocalizations developed so as to express many different emotions, but — due to articulatory problems and to the general difficulty of inventing words — they were closer to song than to speech. Language as we know it emerged only later. Thus, "a primeval song-language is the transition that leads the first societies from instinctive cries to language and reflection" (Thomas 1995: 72). It follows that tone languages like Chinese are thought to have remained close to the original type of language. (For a general discussion of Condillac's theory of language, see Itkonen 1991: 272-274).

In his *Essai sur l'origine des langues,* Rousseau too assumes an early coexistence of voice and gesture, but he sees their respective roles quite differently. For him, gesture expresses what is private and rational, i.e. needs (and later, ideas), whereas voice expresses what is common to many and emotive, i.e. passions. In the beginning, voice was predominant: "the first languages were songlike and passionate". Since then, however, languages have become corrupted — Northern languages more so than Southern or Oriental languages — by the gradual increase of the rational element (cf. Thomas 1995: chap. 4).

Let us return to the present day. The modern authors mentioned above are united in rejecting the well-known scenario according to which language originated as a system of monosyllabic grunt-like sounds that had the (referential) function of pointing at things. Instead, they propose a scenario according to which referential meaning was slowly grafted upon nearly autonomous melodious sound. This is an ironic vindication of Chomsky's overall position. All along, he has been arguing for 'autonomous syntax', or the view that in human language, *as it exists today*, form is primary with respect to meaning. This view is patently false. The most obvious arguments against it are brought together in Itkonen (1996: 483-486). But now, as a result of the foregoing excursus into theories about the evolution of language, we see that (relative) autonomy of syntax may well have been true of that type of vocalization which preceded language proper.

References

Clark, Herbert (1996) *Using Language*. Cambridge: Cambridge University Press.

Davies, Martin (1994) *Musical Meaning and Expression*. Ithaca: Cornell University Press.

Itkonen, Esa (1978) *Grammatical Theory and Metascience*. Amsterdam: Benjamins.

Itkonen, Esa (1983) *Causality in Linguistic Theory*. London: Croom Helm.

Itkonen, Esa (1991) *Universal History of Linguistics: India, China, Arabia, Europe*. Amsterdam/Philadelphia: Benjamins.

Itkonen, Esa (1994) Iconicity, Analogy, and Universal Grammar. *Journal of Pragmatics* 22: 37-53.

Itkonen, Esa (1996) Concerning the Generative Paradigm. *Journal of Pragmatics* 25: 471-501.

Itkonen, Esa (1997) The Social Ontology of Linguistic Meaning. *SKY: The Yearbook of the Linguistic Association of Finland*: 49-80.

Itkonen, Esa & Jussi Haukioja (1997) A Rehabilitation of Analogy in Syntax (and elsewhere). In András Kertész (ed.), *Metalinguistik im Wandel*, pp. 131-177. Frankfurt a/M: Peter Lang.

Jackendoff, Ray (1987) *Consciousness and the Computational Mind*. Cambridge MA: MIT Press.

Jackendoff, Ray (1992) *Languages of the Mind*. Cambridge MA: MIT Press.

Jespersen, Otto (1922) *Language: Its Nature, Development, and Origin*. Oxford.

Kenny, Anthony (1973) *Wittgenstein*. Harmondsworth: Penguin Books.

Klima, Edvard & Bellugi, Ursula (1976) Poetry and Song in a Language without Sound. *Cognition* 4: 45-97.

Koestler, Arthur (1967) [1964] *The Act of Creation*. New York: Dell.

Langer, Susan (1942) *Philosophy in a New Key*. New York: Scribner.

Sapir, Edward (1921) *Language*. New York: Harcourt.

Staal, Frits (1989) *Rules without Meaning*. New York: Lang.

Thomas, Downing (1995) *Music and the Origins of Language*. Cambridge: Cambridge University Press.

Tye, Michael (1991) *The Imagery Debate*. Cambridge MA: MIT Press.

Wallin, Nils (1991) *Biomusicology*. Stuyvesant NY: Pendragon Press.

Esa Itkonen
Department of Linguistics
Henrikinkatu 4a
FIN-20014 University of Turku
Finland
E-mail: eitkonen@utu.fi