

Per Aage Brandt

**Music and How We Became Human —
a View from Cognitive Semiotics**

I. Introduction

Our species, *H. sapiens*, was biologically stable and physiologically modern 160.000 years ago (Singer, 2003). When glaciation stopped, 150.000 years later, agriculture, writing, and History — cultural life based on a symbolically represented shared past — emerged. Somewhere in the middle of this long interval, maybe around 50.000 years ago, humans apparently began to 'make sense' together — to symbolize, paint, speak, form kinship systems that held communities together — and according to the scenario I wish to propose, maybe first of all, they made music. It is commonly estimated that in the Upper Paleolithic, during the Würm glaciation, humans equipped with Aurignacian technology¹ and cooking by fire began to paint in caves and to make music by beating on resonant objects, blowing in hollowed objects (flutes), striking tunable stalactites, etc., and dancing.

There are several ways to arrive at the hypothesis that musical practice may have preceded symbolic, or intentionally semiotic, practices of modern humans. Here follow some such arguments.

1) Death and Danger:

Memory-based feelings, such as those related to the collective commemoration of the dead and to different ritual forms of imaginary communication with remembered persons, and hence the cult of ancestors, the belief in their existence as spirits and ghosts, etc., and experiences implied in the convocation of these spirits, are probably the ancestors of modern 'gods'. Especially in situations of collective crisis, these immaterial beings are called upon through ceremonial performances — activating the genres of human sensitivity and activity that we now call religious— and such events are, in all known cultural communities, linked to musical performances. The shared experience of rhythm and tonal articulation of the human voice (into

¹ Cf. Encyclopedia Universalis (1985 sq.), art. Paléolithique, Préhistoire.

stable tones and intervals²) and of the sounds of melodic and rhythmic instruments universally affect our embodied minds by creating what we could call 'non-pragmatic states', i.e. states of non-functionality, of contemplation, exaltation or even trance, that are regularly expected and presupposed in situations of sacredness: celebration, commemoration or invocation. Collective musical practices also form the aesthetic framing of many trivially pragmatic (work-related) forms of communication, such as the institutional genres of functional verbal communication, which still often entail occasional hymnic singing, performative chanting, ceremonial choreography and gestural coordination (politeness). School assemblies, parades, even contemporary tv news programmes, would be good examples of quasi-pragmatic uses of music.

2) Traces of music in language:

In all known languages, regular intonation patterns connect lexical items and syntactic constructions. From syllabic quantity, stress, and tone, to clause melodies and syntactic stress, and from there to global intonational profiles marking utterance modes and discourse genres (narrative vs. argumentative, exhortative, imperative etc.), linguistic expressivity thus includes and integrates levels of phrasing in a musical sense. This fact is well-known in modern general phonetics.³ Furthermore, dialogical rhythms of turn-taking and attunement to emotionally determined styles of legato, staccato, or rubato phrasing in different tempi are important for the proper use of language in conversation and in the performance of speech acts. These constitutive 'suprasegmental' structures may — although it is of course hard to find conclusive evidence for this speculative hypothesis — be a sedimentation of antecedent and still active underlying forms of musical expressivity. Clause embedding (such as the insertion of completive, relative or adverbial phrases in a matrix sentence) is freely phrasable in oral expression by changes in tone and tempo: in ordinary speech, we spontaneously 'sing' the overall structure of our grammatical sentences in accordance with the intersubjective circumstances and our purpose. Remarkably, there is no well-established

² The discretization that transforms an original glissando into a series of distinct tonal steps is crucial to the change from shouting to chanting and singing.

³ Intonation profiles universally distinguish imperative, interrogative, affirmative, and affective modes of utterance meaning.

theory of the origin of this phenomenon. Nevertheless, we need only pay attention to the role of playful singing and rhyming in infant and toddler language acquisition from early babbling up to the multi-clause stage to get a strong demonstration of the formative force of expressive musicality.

3) Language in music:

The transformation of sentences into verses is universally manifested and universally understood as a poetic device. Poetry exists in all known cultures as an aesthetic genre of oral expression built on significant integration of text and instrumental music, by which in principle some sense of music frames the text: poetry is chanted, sung, or solemnly recited, often on a background of accompanying music, maybe primordially by a string instrument. Even when the music seems to vanish into a silent metric pattern underlying verse lines, and thus leaving the unaccompanied verse as such, as a formal framework for the poetic genre of signification, (e. g. in the academic poetry of the last five centuries of Western culture), this rhythmic framing or integration remains phenomenologically constitutive of the poetic phenomenon as such. In poetry, language and music are integrated by a surprisingly smooth unilateral mapping from the former to the latter. – If language already 'contains' and builds upon music, this operation of transposition is of course easier to explain. (Turner & Pöppel 1983, 1988; also Miall & Dissanayake 2003). Still, language is a distinct structural property of the socialized human mind. It acquires a triple compositionality of its own — phonetic, syntactic, and semantic components that allow us to think and share ideas of absent, past, distant things — and we do not currently know in detail how our mental and neural architecture has shaped its relation to music. We do not yet know if music and language have evolved independently, or if language could have evolved without music; it remains nevertheless a plausible hypothesis that language emerged as 'embedded' in music – that poetry preceded prose. To our counterfactual imagination, it appears that these two semi-automatic communication systems, language and music, would both be reduced to functional signalling systems with limited referential or explicit narrative power if they had been and stayed mutually unconnected. There is — I believe — something in music, or musicality, that language structurally needs in order to be symbolic in the technical sense, that is, in order to be able to

intentionally refer to states of affairs outside of the deictic 'here and now' of persons in communication. This 'something' includes in particular the invocational effect of rhythm in expressive movement.

In the following sections, I will present a series of more specific ideas on the role of music in the constituting of humans as a 'symbolic species' (Deacon, 1997).

II. An indispensable emotional background

There is overwhelming evidence of a fundamental, stable, and primordial connection between music and feelings, and in particular emotional states related to the inter-human affective state we call love. Linguistico-musical compositions in the world literature of scores and texts of songs, lieder, hymns, dramatic works, ballads, operas, and language-related musical creations in general show a constant semantic preference for this affective category as a thematic focus. Poetry in world literature is predominantly 'about' this particular theme and the affective state of love. Such a semantic binding to a specific preferential domain of content calls for a semiotic reflection. There must be a very strong connection between this realm of affective states between persons and 'musicality'. My rather unromantic suggestion is the following.

Once the technology of tools and weapons allowed males and females of our species to extend their respective territories of operation, namely male long-distance hunting and female stationary roaming and especially fishing, something like what we call 'couples', or adult parenthood partners, must have endured longer periods of separation. Fishing allows stationary living habits and thus favors stationary nursing. The fine motor digital skills of females, manifested in the production of adornments and fishing tools, could also be developed during the same period of early symbolic constitution. (Cleyet-Merle, 1990). The concept of parenthood, family relations, stable partnership, the first modern notion of a 'loving couple', presupposes a capacity to recall and recognize the (significant) Other, to identify the beloved's face and person and to associate these permanently with a given

proper name. Names in this sense⁴ are not used for referring to trivial artefacts or objects and animals in general, but primarily to refer to persons (in this modern sense), and hence to personal belongings and territories. But proper names have other absolutely decisive semiotic qualities. They make it possible to designate only the numerical identity of an individual, thus to signify the singularity of a given individual entity, not the qualitative properties of that entity — precisely what we do when naming persons.⁵ Furthermore, whether the entity is a person or not, once the principle is installed, the signified singularity of an item makes it possible to 'cognize' it as an abstract ontological entity, a 'countable' being of some kind, perceived with a numerical, i.e. radically individual, self-identity, stable through time, precisely like a 'love for a lifetime' addressing the (same) 'one and only' person. This emotional binding to 'one-ness' is a cognitive capacity that in other species appear to be much more vaguely present, and which humans in certain psychopathological states again tend to lose.

Nostalgic songs expressing the longing for an absent beloved person are probably a basic genre of all texted music. The name of the beloved is a quasi-obligatory part of such songs. A contemporary jazz songbook will include: I loves you Porgy; Dindi; Stella by Starlight; Michelle; My Funny Valentine; Sweet Lorraine... Then nostalgic or grieving songs recalling deceased loved persons in general will follow the same pattern and will as vividly evoke the spiritual presence of the person thus designated. Names are small phonetic songs in themselves, and the melody of a name song can easily identify a person (a thematic principle later exploited in opera and cinema). When we vocally call each other at a distance, the melodic aspect of the sound sequence is particularly efficient. (Something like the note series C-A-A-F is often heard as the melody for calling Se-bas-ti-an, with reduced versions like A-A-F for Jo-na-than, and just A-F for John-ny...). The local point I wish to make here is that proper names should be understood from the point of view of music and personhood: these nominal entities are arbitrary, emphatically conventional, symbolic signs established by performative rituals, and

⁴ Proper names and common nouns are linguistically and semiotically distinct phenomena. Nouns pertain to the natural mental process of categorization, whereas proper names are grounded in speech acts and possessive intersubjective relations.

⁵ The philosophical distinction between numerical and qualitative identity is not common knowledge in contemporary culture. Sameness refers either to an individual's continuous

basically 'mean' or refer to the affect (love) that first made an individual into a person, a subject inscribed in kinship relations and recognized as a singular and personalized being.⁶ Names are of course intimately related to parental feelings, to the procedures of 'giving' names, analogous to the idea of 'giving' life, and especially to the existence of a universal practice of voiced interaction between infant and parents (Trevarthen and Malloch, 2002; see Ellen Dissanayake, Chapter 2).

However, this love-borne 'nominalism' and personalized orientation in music and poetry, by which music inherently seems to 'think about' love, and love seems to 'think about' music, needs in its turn a grounding in other semiotic factors and circumstances, such as those we will consider in the following sections.

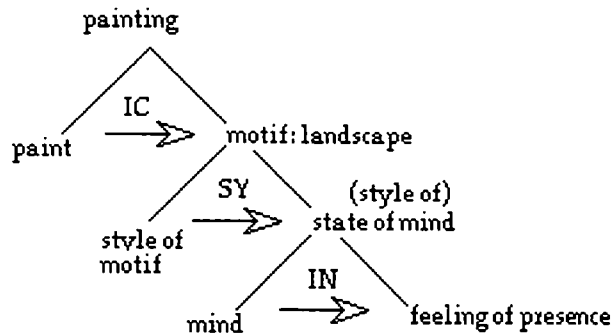
III. Homunculus in the artistic and musical sign

The global musical sign constitutes in itself an important prerequisite for its emotional use. Let me explain this semiotic phenomenon after first presenting a pictorial analogy. A painting, for example a landscape, offers primarily an iconic relation between a canvas framing a complex multitude of graphic and chromatic events appearing on the painted surface, and a framed view of the 'depicted' landscape, as seen from a window, or from some other limited vantage point. The landscape in question may be some real place whose name appears in the title of the painting; it may thus be a representation of an existing geographic locality; or it may be a pure invention by the painter. To the observer, it shows a fragment of a generically possible 'world', in that it offers a supposedly intentional glimpse into such a 'world' or place, and in such a way that the properties of the glimpse or part would illustrate the general character of the whole it refers to. The part symbolizes an underlying, more general whole, which it 'stands for'. So the initial icon gives rise to an intentional act of symbolization, and the landscape painting is now a symbol of a character, style or atmosphere, or state of mind, of some possible spatial habitat. Finally, since the painting in front of us addresses our attention without further specification, we 'read it' as an unspecified, existential index:

existence through time or to a property shared by several individuals. I am 'me' by numerical identity, and I am such and such by qualitative identity.

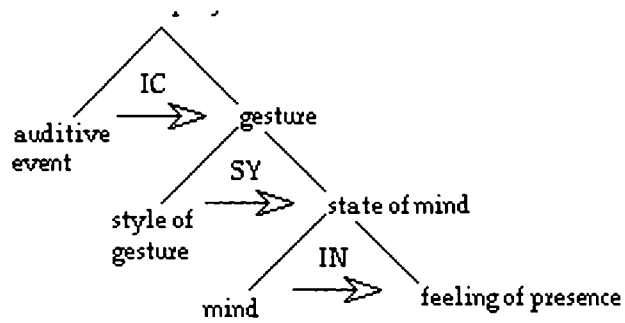
⁶ Notice that it is still the case that we find animals carrying a proper name more difficult to eat than anonymous creatures.

a human mind was there and as a materialized symbolizer still is here, now, with us, through the presence of his work.⁷ A painting is thus a cascade of sign functions, IC → SY → IN, where the content of its icon (IC) is again a sign, namely a symbol (SY) whose content is a sign, namely the index (IN) whose content is the presence of the 'ghost' (spirit) of the artist. In the example we may model such a triple sign as follows (Figure 1):



The triple sign produced by a musical performance or experience can be compared to what happens in pictorial iconicity. 1) The rhythmic and melodic gesture will suggest a body making that gesture. In this sense, the auditive form iconically means (signifies) that bodily gesture, even if it is not actually shown but only 'played'. 2) Furthermore, the idea of the gesturing body will symbolize, symbolically mean and signify, a person in a corresponding general state of mind. 3) Since this very abstract affective meaning, or symbolic content of the iconic sign, occurs in the very moment of auditive musical perception, it will ultimately and indexically yield a feeling of the presence of a 'ghost' or 'spirit', or 'avatar' (or whatever we might need to call it) of that person, to the persons sharing the musical experience. So, in the case of a musical event, we would have a corresponding sign cascade (Figure 2):

⁷ Symbolization always yields metonymic presence of the symbolizer. Inversely, it may be true to say that the semiotics of metonymy always involves an act of symbolization of some sort.



The cascade format of semiotic meaning processing is rather clear and obvious to most people in these and related cases of art; however, the semiotic cascade may also be cognitively active in other forms of communication by explicitly expressive signs, whether arbitrarily coded or not, such as facial expressions, theatrical gestures of politeness, pragmatic signposts, signboards, etc. The particular interest humans take in the artistic cascade, however, is undoubtedly due to the forceful feeling created by a particularly elaborated iconic stance in art, by which the symbolic function is built into the content of the icon and therefore made immanent and disembodied, so that the symbolized emotional state of mind does not carry the signature of the performer but will instead stay an immanent semantic property of the artistic piece of work. The participants will be able to feel, sense or accept the emotion of the state of mind in question without 'being in it'. The subject of the mind whose presence is felt by the participants is what I suggest to call a 'homunculus', an imaginary persona or 'virtual other' experienced as immanent in the work of art.

When art is associated with cultural, institutional, and discursive practices of different kinds, including religion, the authority of a voice experienced as emanating from an artistic expression will then be associated with the abstract homunculus – whose disembodied status will endow it with a particular symbolic force, maybe the dynamic effect we call 'sacredness'.

In the evolution of cultural practices, in the first place, I would therefore claim that the necessary presence of such authority-yielding symbolic forces – especially in the execution of performative acts and rituals – stems from the semiotic homunculus. Music generates sacredness.

Furthermore, it is probable that the visually based cascades appear in evolution subsequent to the auditory based cascades. It is only possible to derive the symbolic meaning from the iconic content to the extent that

different modes of representation can be perceived as 'styles' or graphically manifested expressive gestures (responsible for strokes, colours, contours, light) characterising variable mental 'styles' or perceptive modes of seeing. Musical rhythms (cf. strokes), soundings (cf. colors), melodic phrasing (cf. contours) by contrast directly inform our bodies of the way to dance in order to unfold their meaning; we immediately grasp the state of their 'homuncular' mental being, or 'spirit'. Meaning, as distinct from the fact of someone who 'means' something by saying it, is homuncular. It transcends its performer.

In so far as this privilege has always been a property of our motor-based perception of temporal events, music may have guided other expressive modalities and eventually language – the voice heard in different forms of enunciation (irony, bathos, imperatives, interrogatives etc.) are indirectly, theatrically linked to the speaker and directly related to this homuncular symbolic force. Implicit narrators in fiction and humour, impersonal bureaucratic formulaicity, juridical textuality etc. all rely on homuncular enunciation. The law 'speaks', or rather chants, and we can sympathize with this authoritative voice or mock it by letting it sound like good or bad music.

There is a structured process in the architecture of the human mind that 'does' semiotic cascades and expressive body codings associated with them, and that represents the virtual, homuncular other in relation to the Ego. Let me briefly and speculatively outline the general semiotic view underlying this line of thinking about cognitive aesthetics and musicology.

IV. Mental architecture and the role of music

The human mind organizes knowledge about the spatial and temporal world, including the body that hosts it. Additionally, but as importantly, it organizes the functional and expressive acts of 'its' individual host as an embodied person in society, namely in a society of persons sharing significant homunculi, while sharing imagery and music. Thus, we 'perceive' but also 'perform'. To account for this double perspective of our subjectivity in theoretical or philosophical terms is highly complicated; current research is far from having an elaborate model at its disposal for orienting its technical and empirical investigation. Nevertheless, there are certain elementary

principles that begin to emerge, allowing us to form a first, minimally ordered view of what the mental brain is doing.

Two dimensions must be distinguished:

1) a 'vertical' dimension in which afferent integration builds up content from 'input', and efferent integration builds up our agentic programmes as 'output'. In this sense, borrowing terms from neuroscience, we could speak of afferent and efferent cognition. (Afferent: bearing or conducting inwards; in neurology: conveying impulses toward the central nervous system; efferent: conducting outward from an organ; conveying impulses to an effector).

2) A 'horizontal' dimension, in which different levels of mental work are articulated, i. e. separated and connected.

a) In afferent cognition, five superimposed levels of distinct and relatively independent conscious meaning production, as a minimum, appear to be operating in parallel: (1) perception which precedes (2) categorization, and conceptual categories which in turn precede situational scenario formation, also called (3) narrative cognition; (4) comparative and reflective recall constitute a fourth level of consciousness, and (5) there is an ultimate level of free-floating imagination and 'off-line' representations, ideas, day-dreams etc. In this order, each level presupposes systematic access to the products of the preceding level.

b) In efferent cognition, the last level of the outward oriented process, which shapes our bodily reactions to the surroundings we perceive, must be closely related to the first level of the afferent process by some sort of bridge, creating a shared level, since specific sensory perceptions (gestalts) can directly and spontaneously trigger certain gestures and reflexes, typically such deictical moves by which we apply volition, positive or negative, to what we sense, in order to better perceive it. Behind this level of deixis and volition, or underlying it, afferent categorization must be connected to efferent object-oriented motor routines by a second bridge (and the bridge on this second level between afference and efference may therefore be related to lexical structure in language). On a still deeper level, efference prepares sequences of acts that express superordinate intentional meanings, connected to afferent situational understanding by a narrative organizer of temporal experiences (related to semio-syntactic structure in language). Underlying this level, our semiotic body finds its affective tonus, or emotional attitude, by

we listen and speak through the same grammatical forms); only in foreign language acquisition do we observe a significant difference in afferent and efferent competence. We are, incidentally, normally better at reading (hearing) than at writing (speaking) a foreign language. This difference is probably due to the role of consciousness of others' actions and expressions in language learning – for many reasons, it is easier to attend to reception (afferent content) than to production (efferent content) if one is oriented to the reception of a message intended by someone. Others are apparently more salient in the afferent than in the efferent line of processing.

If the architectural hypothesis here presented is solid, then music is essentially both a matter of auditory perception and of deep, abstract ideation, an ideation that originates with the impulse to move, i.e. with action. Whereas auditory events are generally perceived to be integrated into multimodal clusters of objectal concepts – since 'things' yield multimodal sensations – musical sounds are perceived as tones, which have rhythmic meaning as beats. We need to ask what the particular principle underlying this truly strange fact could be. Of course, the 'strange fact' is comparable to what happens in visual art and pictorial and graphic iconism in general: the visual mode is kept separated from other possible sensory gestalts offered by the source of perception. Otherwise, there would be no 'image'. The auditory percept is thus carried through all standard instances – categorization, narrativization, reflection, imagination – without being absorbed by contextual meaning, and is then interpreted as an event manifesting the spiritual presence of some being! We may explain this symbolico-aesthetic miracle simply by stressing the basic fact that musical sound is perceived as an intentional gesture, i.e. as a 'symptom' of someone moving in a particular expressive way. Since it is immediately understood as an intentional expression, attention is drawn toward the category: Other Person's Conscious Doing. And since musical sounds are produced by persons producing them knowingly, the actual Other Person playing is conceived of as particularly self-conscious, so that there are inherently three intentional processes going on at the same time: a listener's conscious attending, a player's conscious attending to what is played, and the consciousness invested in the music that the player attends to while playing it! The latter intentional instance is precisely what the listener foregrounds. Not the player's auto-controlling but

the musical flow the player intends to control, and which is thus objectified during its production as an autonomous instance: the meaning of what is played.

This rather tricky phenomenological analysis may be fundamental to the general understanding of our topic, so I will rephrase it a couple of times. To play or paint something (instead of just performing intransitively) is to embody and inhabit this something and to experience it as a pre-existing efference that the actual efference emulates or reactivates. It is very particularly this pre-existing 'intentionality' that the musical or pictorial experiencers focus on – beyond the performer's own efference. The triple subjectivity generically built into the process stems from the performer's normative project, one could say. Since the performer creates 'something' and thereby could either fail or succeed to give birth to it, according as the meaning immanent in the 'something' is saved by the performance, the feeling of a precarious, fragile transcendent intentionality accompanies quite naturally the aesthetic display.

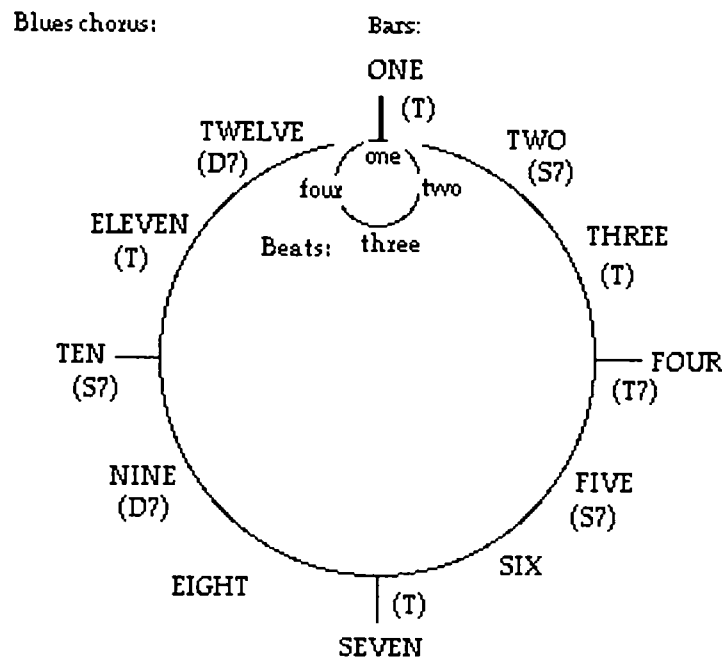
Every act of symbolization is a normative performance project in this sense and therefore entails the feeling of transcendence that we encounter here, and which is more directly and clearly present in the musical here-and-now experience than in any other circumstance. Symbolization may thus be derived from the primordial musical practice of humans, as suggested above.

There is still, however, a constitutive aspect of symbolicity that needs to be elucidated: how did we manage to isolate symbols as discrete single signifiers and then to conceive of their combinations as formulaic sequences? Where could this discretization and this idea of concatenation have entered human cognition? Again, music may have been a structural source of these formal cognitive inventions, as I will briefly try to show in the following, concluding section.

V. Names and numbers: from metric and rhythmic time to calendar time

In non-written music worldwide, as in most academic score-based music, the rhythmic organization that musicians, singers, dancers anchor their playing and performing in, consists basically of some finite temporal units that can be described as recursive 'measures' or 'bars' comprising a short sequence of regular pulses, or 'beats'. These bars form a shared reference for the

performers and allow them to synchronize their expressions. The finiteness of the bar makes it possible to conceptualize a temporal flow as a highly structured recursive process of nested metric cycles. The encompassing multi-bar units (cf. the 'choruses' in jazz) are normally related to melodic wholes, and there are further compositional, multi-melodic wholes, united or separated by specific scales and harmonic preferences. Here is, just for the sake of demonstration, a 12-bar blues schema with T, S, D = tonic, subdominant, dominant (Figure 4):



A construction like this one is only possible because the beats of the bars are numbered (named), so that a musician could count: one-two-three-four, two-two-three-four, three-two-three-four, etc. – the first one referring both to the beat and to the bar: a double present, so to speak. This is already in itself and as such a numerical system: it is tetrachic, comparable to the decimal or the binary systems, and the possibility of identifying a unit by at least two recursive parameters is what makes the unit as symbolic as a person's name, including first name and family name. Once we are able to name a beat, within a closed list of possible names, we can conceptualize the temporal

moment as a 'place' in time: a recurrent place as something to return to, something immaterial which is still there "as time goes by", so that different persons' presence there in the future can coincide or significantly not coincide. Planification becomes possible. The calendar is born. The elementary miracle is of course that the place will be there whether or not someone pays it a visit: the beat, and equally the bar, exist even if they are empty! An empty – that is, unmarked, unplayed – beat is an auditory event that we do not hear, it is an acoustic 'ghost', one could say. It exists plainly, numerically; and, I claim, this is how plain natural numbers might have come into existence – as beats to fill or leave unfilled. A named beat is a numerator with an unfilled, pronominal denominator. The embodied origin of mathematics might thus be the nested cyclicity of musical rhythm.⁸

Note that the metric underpinning of poetic rhythm – beyond the quibble of feet, tones, accents, and quantity in culturally distinct poetics – is exactly the same beat-based temporal cognition. Here is a stanza by Robert Burns (1791: On Mary, Queen of Scots):

O! Soon, to me, may Summer suns
 Nae mair light up the morn!
 Nae mair, to me, the Autumn winds
 Wave o'er the yellow corn!
 And in the narrow house o' death
 Let Winter round me rave;
 And the next flow'rs, that deck the Spring,
 Bloom on my peaceful grave!

Four beats organize each verse as a bar:

1	2	3	4
O! Soon, to me, may Summer suns			
Nae mair light up the morn! – [4]			

Note the empty fourth beat in line 2. The syntactic accentuation would at the close oppose the realization of these rhythmic beats by strongly stressed syllables:

⁸ Lakoff and Nunez (2000) prefer to believe that numbers are grounded in "subitizing" our fingers (p. 52).

An **the** next **flow**'rs ...

Bloom **on** my ...

Here, the linguistically unaccentuated morphemes would be grotesquely overstressed, if their stress were to follow the four-beat rhythm; instead, they are to be pronounced in a slightly slower tempo and with an artificially equalized half-stressed weight, a counteraccental solution that yields a perceptible poetic effect.

Let me present one more example, a famous Japanese haiku by Matsuo Bashô (1644–94):

Furu ike ya	[an old pond and]
Kawazu tobikomu	[a frog jumps]
Mizu no oto ⁹	[water's sound]

It is also known that the verses of a haiku have 5 + 7 + 5 syllables. The stressed voicing of these lines, however, uses a four-beat measure:

1	2	3	4
Furu ike ya	[4]		
Kawazu tobikomu			
Mizu no oto	[4]		

The result is that the final void [4] – the empty beat following *oto* – becomes the temporal place of the splashing beat. A poetic trick consisting in animating the void, or rather: semanticizing the pure temporal slot.

We know that music has always and transculturally been associated with the hours of the day and the night; in fact, the notion of hour and day is due to the same nested cyclicity as the musical metric itself. The names of hours are mostly numerical, and this is often as well the case of days (cf. Portuguese weekdays: *segunda-feira*, *terça-feira*, *quarta-feira*, *quinta-feira*, *sexta-feira*). Socio-cultural conceptualizations of time are throughout isomorphic with time's musical form of schematization. It is evident that calendars – using names of divine entities as 'numbers' – are built out of

⁹ In (ed.) Stryk and Ikemoto 1977, the translation is (p. 91): Old pond, / leap-splash – / a frog. Cf. also <http://www.teeweg.de/de/literatur/basho/furuikeya.htm>

exactly the same symbolic substance. Symbolization springs from temporal cognition, I conclude.

Let me add a last remark on tones. The discretization of tonal sounds, already specified as tones, not noises, by their formants (overtones), and the melodic combination of tones of different pitch, as produced by musical instruments perceived as analogous to the vocalizing human voice, probably occurred when they were connected to beats. A tone manifesting a beat calls for subsequent tones representing other beats of the same cycle or multi-cycle (cf. the blues cogwheel in Fig. 4). So the length of the tone comes to refer to the beats of the bar as a metric, quantitative scale; as there is no cognitive continuity or gradual transition from one beat to the next, the tonal signifier of the beat will be cognized as a discontinuous, discrete, sounding event with a determinable onset and followed by a new onset of a tone, same or distinct, or by a pause (a void beat). Since the rhythmic organization is serial, finite, and cyclic, this alliance of tone and beat leads to the invention and stabilization of finite scales – series of notes separated by stable intervals and united by their affinity as sets of elements that combine syntactically into cognitively clear melodic forms. As soon as a note is integrated in a scale, it acquires a name (e.g.: do - re - mi - ..., c - d - e - ... in modern terms). Scales are sometimes associated with affective moods and social situations, in such a way that a musical culture will dispose of different scales felt as appropriate for correspondingly different moods and situations. (This is explicit in the genres of flamenco music). In a sense, these scale systems are psychological and sociological 'theories' in themselves. They interpret significant moments of shared human time, with universal emotional appeals as well as conventions of acceptance.

Discretization (tones are discrete units, not glissandos) and finitization (beats are members of finite recursive series, not elements of an unending train) are thus basic aspects of the genesis of symbolic expressions. When the human voice finally stabilizes the sets of linguistic sounds we call phonemes, it does so within a phenomenology of syllables, but on this phonotactic level – more easily experienced than single consonants and vowels – discretization and, to a certain extent, finitization likewise takes place. The syllabic phenomenon, including the naturalness with which we articulate words by dividing them into syllabic sequences, could be an effect of the musical

binding of tonal sound and beat. Syntactic phrase formation would be an additional melodic superstructure. The reproducibility of melodic phrases and their easy interpersonal transmission, due to the particular refinement of our auditory memory for effects of action, makes them appropriate for interpersonal monitoring of attention to situations; and since melodic integration does not eliminate the discontinuity of its syllabico-lexical components, the tension between separate words and integrative clauses, as between tones and melody (a dynamic principle exploited in thematic variation) generates what we call grammar. Grammar is not a system but rather a constant crisis: words do not dissolve into phrases or clauses, but instead tend toward discreteness and autonomy. Oral phrases are thus often completed by gesture and intonation rather than explicit wording. We can start a sentence with explicit wording, and then go on with a "nanana and nanana..." that every hearer will understand. The superordinate intonational utterance profile would eventually represent a supplementary expressive unification or homogeneization of discrete units, rooted in affect and rhythm, as we have seen.

Musicality of action and consciousness is possibly even the factor that unified the lexical (object-oriented) and the syntactic (propositional and evaluative, subject-oriented) components of language and thereby created the very logos of our species. The first manifestations of language would therefore have been what we now call poetry. Maybe the cave paintings in the deep, acoustically rewarding halls, where no signs of household are found, were the scores of the recitals and musical performances that shaped human culture. Their superimpositions of figures may even have corresponded to melodic superimpositions in baroque fugues. Synaesthesia is now understood to be common in the perception of beauty, as it is, indeed in any active experience in a richly stimulating world.¹⁰ Maybe Cromagnon humanity emerged as a bouquet of baroque cultures, first using petrophones (resounding stones) or stalagmites, flutes, and drums, then bowed strings and (animal) horns, to create and animate the cognitive and emotional architectures that eventually grounded imagination and rationality.

¹⁰ Cf. Ramachandran and Hubbard 2001 for a particularly interesting study of and reflection on synaesthesia, art and language.

In contemporary and future research, musicology and many different forms of cognitive and semiotic studies may be collaborating along the lines of these or similar evolutionary scenarios, hopefully elaborating both imaginative hypotheses and finding still more empirical arguments in favour of a coherent reconstruction of the origins of human symbolization; what is already overwhelmingly probable is that symbolization is grounded in temporal cognition, and that human conceptualization of time is grounded in music.

*

Bibliography

- Brandt, Per Aage, 2004, Spaces, Domains, and meaning. Essays in Cognitive Semiotics, European Semiotics No. 4, Bern: Peter Lang
- Cleyet-Merle, Jean-Jacques, 1990, La préhistoire de la pêche, Paris: Ed. Errance
- Deacon, Terrence, 1997, The Symbolic Species. The Co-evolution of Language and the Brain, New York: Norton
- Encyclopædia Universalis, 1999, Dictionnaire de la Préhistoire, Paris: Albin Michel
- Lakoff, George and Núñez, Rafael E., 2000, Where Mathematics comes from. How the embodied Mind brings Mathematics into Being, New York: Basic Books
- Miall, D. S. and Dissanayake, E. , 2003, "The poetics of babytalk". Human Nature, 14(4): 337-364.
- Ramachandran, V. S. and Hubbard, E. M., 2001, "Synaesthesia – A window into perception, thought and language", Journal of Consciousness Studies, Vol. 8 No. 12, Imprint Academic, pp. 3–34
- Singer, Chris, "Out of Ethiopia", Nature 12 June 2003, pp. 692-693.
- Stryk, Lucien and Ikemoto, Takashi, The Penguin Book of Zen Poetry, Harmondsworth: Penguin Books Ltd.
- Trevarthen, Colwyn and Malloch, S., 2002, "Musicality and music before three: Human vitality and invention shared with pride", Zero to Three, Vol. 23, No. 1, pp. 10-18.

Turner, F. and Pöppel, E. ,1983, "The neural lyre: Poetic meter, the brain and time", Poetry. In (ed. R. S. Gwynn), 1999, New Expansive Poetry, Ashland: Story Line Press.

*