ANALOGY AND CONVERGENCE BETWEEN THE COGNITIVE PROCESSING OF TEXTUAL MEANING AND MUSIC: THEORETICAL CONSIDERATIONS OF A SONG AS A COMPLEX SEMIOTIC OBJECT

ALICE KARBANOVA
(Masaryk University, Brno, Czech Republic)

Abstract

This article is concerned with the relationship between language processing and music processing at the semantic level. It is being investigated whether language and music compete for processing resources during the perception and interpreting of a song. Since music, just like language, represents a uniquely human and universal feature that challenges almost all of the components of human cognition, it constitutes one of the most prominent tools of exploring the human cognitive processes and scholars from various disciplines grant the music-language research an ever growing importance (Fitch, 2006). A song, stimulus combining both speech and music, represents possibly the most natural setting to compare music and language. In the past decades, this topic has gained a significant scientific interest and this paper draws on the latest findings from cognitive sciences and neurosciences. Its purpose is to provide an overall insight into the processes involved in perceiving and understanding a complex semiotic object and to answer the question whether or not the array of musical representations becomes predominant in the song listening. In this paper, the evidence bearing on a selection of experimental studies on the music-language interface is reviewed. Related issues, such as the question of domain-specificity and overlapping representations are addressed in an attempt to shed light on the processing of semantics of a song. The discussed evidence suggesting common neural processing networks for music and language as well as independent components of its processing.

Keywords: Song Perception, Musical Semantics, Cognitive Functions, Modules, Specificity

1. Introduction

Music and language seem to draw on similar cognitive processes. I am going to discuss commonalities in processing of meaning in the two domains in an intermediary object, namely a song, material that is hardly ever addressed experimentally and that imposes great challenges on the perceiver as well as on the scientist who is trying to decipher the processes underlying its understanding. Most of the experimental studies focusing mainly on memory and recognition, the semantics remain largely unexplored, one possible reason for that being the uneasy way to address it experimentally. Research of music has proven important for an encompassing understanding of human cognitive processes, since it is thought to be only through an intricate interplay of cognitive processes of manipulating symbolic forms that music acquires its meaning (Molino, 2009, p. 307). It can indeed help us investigate the nature of these processes since specific aspects of musical form may result from constraints imposed by the vertebrate nervous system (Fitch, 2006, p. 184). Our ability to understand is constrained by the natural limits of perception as well as cognitive limits. Since tonal music is founded on the natural principles of human cognition (Gribenski, 2005, p. 19), the recent research is aiming at the very principles of human cognitive faculties. Note that in the present study the word 'meaning' is used in a rather loose sense, assenting to Nöth's statement 'The meaning of *meaning* is a semiotic labyrinth both on theoretical and on terminological grounds' (Nöth, 1995, p. 92).

Addressing specifically song as a middleman of music and language I am going to point out the most significant findings suggesting the power of music to blur the linguistic meaning. The use of ecological experimental material such as song seems furthermore preferable and justified, since the use of artworks can provide insights into general brain functioning (Wassiliwizky & Menninghaus, 2021). This need is also well reflected in the emergence of new fields of research such as the empirical aesthetics that elucidate the general cognitive, emotional, and neurophysiological foundations underlying aesthetic experiences, including their structural and temporal organisation in the brain, and their subjective feeling. Such material represents nevertheless a challenge, given that the association of language and music must be approached as a whole, not as a summation of two separate systems (Vallespir, 2010, p. 37). Study of song, combining poetic expression and music, has called for researchers from different fields. Its processing is a fundamentally interdisciplinary endeavour that has given rise to new disciplines insisting on empirical approaches to the study of literature reception, such as cognitive poetics that sets out to elucidate how poetic language and literary form are shaped and constrained by human cognitive processes (Rasse et al., 2020, p. 6). Since the 1990's the parallel between the two domains is made based on the cognitive mechanisms brought into play in their understanding (Vallespir, 2010).

The research on the music-language interface can advance our understanding of the neural representation of semantics in general and might ultimately enhance our understanding of the meaning of 'meaning' at large (Fitch & Gingras, 2011, p. 90). Investigation of neural correlates of music has began to broaden our understanding of how human brain processes meaning across domains. I am going to briefly summarise results from selected studies that are of interest to the interpretation of songs.

2. Song, Integration of Music and Language

Understanding of song is a complex hermeneutic undertake. The cognitive relationship between lyrics and tune in song is currently under debate and the central issue is whether they are represented as separate components or processed in integration (Sammler et al., 2010, p. 3572). It seems intuitive that music influences the understanding of lyrics, it has nevertheless proven extremely challenging, if not impossible, to address this question experimentally. In order to tackle the problem, one of the possible roads seems to be approaching the interpretation of lyrics as a message, or an act of communication that is constantly influenced by the external factors, the musical accompaniment in this case.

Since language is just one part of communication (Culpeper & Semino, 2002, p. 321), one can think of the musical accompaniment as an additional contextual information. Although music lacks propositional content, it can be meaningful on the pragmatic level and provide the contextual information to semantic structure (Patel, 2008, p. 327). Furthermore, the fact that music might express ideas in a more compelling, although less specific manner than language (Limb, 2006, p. 3), might make it an efficient tool to specify the equivocal poetic expression, the latter being by definition ambiguous.

Music is generally considered an auditory stimulus, but it has been suggested that perceptual and cognitive representation of music can involve non-auditory (e.g. kinaesthetic) information (Hubbart, 2019, p. 521). Moreover, music might be able to induce an image of a virtual environment through the sense of motion and fictional movements and gestures (Patel, 2008, p. 319). In the case of a song, these abstract mental representations of the movement may be contrasting to the text, which might have an overriding effect on the perception. By the same token, Patel (2008) argues that the ability to evoke a sense of motion and to generate the tendency to synchronise to a musical beat appears to be a uniquely human response to music. Besides, it is not just the neurophysiologically induced self-motion, on top of that music is capable of triggering a sense of external objects moving in relation to the self.

It has been alleged that the musical accompaniment exerts influence upon the text in the way that it points out, highlights, and alters certain aspects of the text through the careful choice of expressive means (Dürr, 2004, p. 17). At the same time, putting a poem into music means intensifying the utterance with the aim to communicate more content. The accompaniment usually tries to follow the prevailing mood of the passage but is able in individual cases to intensify or question the meaning of the text. Furthermore, musical structures form a scene on which the described situation takes place (Dürr, 2004, p. 251), a fact that is in line with the current opinions in understanding on narratives, namely by subconsciously imaging and simulating the situation being described (Feldman & Narayanan, 2004). By the same token, suggestion has been made that music can imply a meaning independently of the semantic load and referential value and provide an auxiliary supplement for the representation (Harmat, 2010). The function of musical accompaniment would therefore be to interpret the text, to reveal its hidden aspects, to bring about unexpected meanings and guide the listener in its interpretation. The melody is supposed to translate what remains hidden in the language and counterbalance its shortcomings (Brogniez & Piret, 2005, p. 11).

Although some researchers defend the view of two rather separated domains (this conviction coming notably from the studies of focal brain damages causing amusia or aphasia, that indicate distinct processing modules for music and speech, e.g. Peretz, 1993; Peretz & Zatorre, 2005, p. 106), there is a growing body of evidence suggesting that the two domains, at least to some extent, involve the same cognitive functions (Lerdahl, 2003, p. 413). It has been suggested that the human brain engages overlapping neural substrates and mechanisms for the processing of both music and language (Besson & Schön, 2001). Since the neural resources underlying their processing appear to be overlapping to a certain degree, it seems plausible to expect transfer effects between music and language (Jentschke et al., 2005, p. 231).

3. Musical Semantics[1]

There is copious evidence for the intricate relationship of syntax processing in language and music, yet a considerable lack of studies focusing on the semantic processing (Koelsch, 2006, p. 518). Nonetheless, in the past decades more attention has been paid to this field and the evidence is mounting that music is not devoid of referential power (Patel, 2008, p. 327), despite the fact that it lacks a rich denotative meaning because its meaning is fundamentally context-dependent (Fitch & Gingras, 2011, p. 90). Although the referential function appears to be absent in music and its mode of signifying is rather metaphoric, that does not mean that music is devoid of meaning (Locatelli & Delpy, 2009, p. 31). Unlike language which denotes specific semantic concepts, music picks out concepts at a coarser grain and the representations it conveys are relatively vague (Kölsch, 2011). At the same time the partial overlap of design features of music with those of language suggests a shared formal core (Fitch, 2006, p. 173). Differences appear to arise from absence of arbitrariness, displacement and duality of patterning, which means that dissimilarity derives precisely from semanticity (Fitch, 2006, p. 173), a fact that might explain the lack of studies in this field. However, language is not a condition sine qua non of human thought and a meaning does not necessarily need to be verbalised. Sebeok (2001) argues that from the phylogenetic and ontogenetic point of view, non-verbal modelling comes always first, and language is from this vantage point regarded only as a secondary modelling system.

Some even grant music a superior power to express certain sensations better, despite the fact that they cannot be translated into words, and consider language as hopelessly limited by its attachment to the signification (Locatelli et Delpy, 2009, p. 30). The inherent vagueness of musical expression represents its very advantage. As Kölsch (2011) puts it, sensations, such as sensori-interoceptive information, action tendencies, or background-feelings have to be reconfigured into words in order to be conceptually grasped and communicated between individuals. Whether these verbalised sensations are shared between individuals is nevertheless dubious even if the exact same word is used to refer to these sensations. Music on the other hand, mediating sensations in their pre-

1 For an exhaustive list of concrete ways music can convey meaning see Kölsch et al. 2004, and Daltrozzo & Schön 2008

linguistic mode of existence, might be much more in accordance with their real essence on the inter-individual level. Such meaning is therefore conveyed prior to its reconfiguration into words. Furthermore, as Fitch & Gingras (2011) affirm, even for language the scientific understanding of meaning remains elusive. We can come near to the meaning of a lexeme but it is illusory to think we can define it once and for all because its meaning essentially depends on the discourse and genre (Cusimano, 2005, p. 35). The context participates fundamentally on the content of a word and the need to approach perception of songs on an interdisciplinary level is therefore urgent.

Following in the footsteps of musical semiotics, the current trend is to conduct empirical experimental studies which have provided valuable insights into the processing of meaning. While there are differences in the semantic organisation of music and language, there is an ever growing agreement on the view that there is also huge overlap with regard to the cognitive processes, and the neural operations, underlying processing of meaning information in music and language. The communication of concepts can be tested experimentally by looking at the effect of the context on the brain activity during the processing of a target stimulus, and a subsequent recording of event related potentials ERPs, which constitutes nowadays a very common means of examination in the musical research, as we shall see below.

4. Pragmatics and Emotional Prosody

It has been pointed out that language users are accustomed to adding contextual information to utterances and that the hearer recovers the intended meaning based on contextual information and inferencing (Patel, 2008, p. 336). Meaning is derived from discourse, operation that involves assuming unstated information, and drawing inferences about what was said. Establishing discourse coherence is therefore central to language understanding. Drawing a path of implication between discourse segments and establishing of coherence relations relies on the sense of cause and effect. It seems that there is no airtight division between the linguistic semantics and pragmatics which is indicative of mutual influences of both domains (Nieuwland & Van Berkum, 2006, p. 1108). It has been suggested that similar cognitive principles are at play in organising the flow of meaning in both language and music. Cross (2011) insists on the participatory nature of musical meaning of which he conceives as a mode of human communication that is homologous with aspects of linguistic interaction, because it is organised around multiple interactional goals that go beyond the transmission and reception of factual information.

Kölsch (2011) suggests that listeners automatically engage social cognition which is reflected by the activation of the cortical Theory-of-Mind network and posterior temporal regions while listening to music. It is believed that the message conveyed by music includes intentions and that deciphering these intentions and the attribution of mental states is one of the components of musical meaning.

Not only do speech and music derive their meaning from the context and discourse, they both possess similar acoustic properties that code emotional expression. The purely semantic content is not the only way an utterance conveys its meaning, far from it. It is essentially constraint by elements external to the propositional content. Notably by the context, as we have already seen, and the prosody. This, in a way musical part of a language utterance enables us to understand the overall message even if it is produced in a foreign language which we do not speak (Dürr, 2004, p. 19). This emotional part of the information, that is much more inter-individually shared and universal, enables the communication and is to a certain extent independent of the codes each individual language uses. Music also constitutes an efficient tool for mood modulation that 'appears as effective as food, drug, and facial expressions to elicit sub-cortically mediated affective responses that result in involuntary changes in physiological and behavioural responses' (Peretz, 2006, p. 23).

One of the ways music conveys meaning is based precisely on the similarity to language prosody of people in a certain emotion state. In spoken language, affective prosody elicits sensational processes in a perceiver that bear resemblance to those that occur in the producer

(Kölsch, 2011, p. 126). This is yet another shared feature of the communication on pragmatic level. Both domains communicate the meaning on the social-intentional dimension. The possibility of carrying meaning through this channel 'derives from physiological and affective constraints that shape human vocal production in ways that render vocal sounds likely to be experienced as indicative of the affective state of the sound producer' (Cross, 2011, p. 118). In addition music can present an appearance of emotion directly, rather than referring as words do, an effect that resembles the contact with the direct expression of an emotion more than hearing its verbal description (Davies, 2011, p. 114).

Humans possess cognitive faculties to attribute mental state to others in a social interaction based on what they hear. The brain has developed mechanisms sensitive for the perception of human voice that are most probably located in the superial temporal gyri and sulci (Grandjean, 2021). There is a special cerebral mechanisms underlying the capacity of the system to process and generate emotional information comprising universals as well as unique traits for each individual language. One of the important functions of emotional prosody is purportedly to capture attention. When the brain detects the emotional information, it orientates the attention towards that stimulus independently of the voluntary attention focus. This phenomenon being essential for the survival.

Martinec (2018) puts forward that rhythmicality of language at the various hierarchical levels differs according to different registers, a fact suggesting that also the rhythm is meaningful, because it is able to categorise the utterance within a certain register. The rhythm of the utterance is constrained in part by language and in part by context that suits the purpose of the moment. In a song, if the musical accompaniment changes the rhythm of the utterance it could shift the linguistic message into another register and therefore impact on the intended meaning. It has been suggested that the musical accompaniment might influence the poetic expression by changing the ordinary organisation of the text. The independent rhythmic part of the accompaniment could in the extreme case make the message unrecognisable (Dürr, 2004, p. 253).

5. Distributed Meaning

One of the views of semantic memory is the distributed-only view which considers widely distributed regions and their connections as the site of the whole semantic network. It is supposed that there is an overlap of concepts having similar semantic significance but not necessarily similar specific attributes, and the function of semantic memory is to generalise across these concepts (Patterson et al., 2007, p. 977). It is also believed that conceptual representations abstract away from modality-specific attributes and higher-order generalisations are necessary for the semantic processing. These findings are consistent with several recent studies demonstrating that the working memory system underlying sentence comprehension is not domain specific (Fedorenko, 2006, as cited in Fedorenko, 2007, p. 6).

Meaning in general is usually multidimensional. It emerges from sign qualities, structural context, idiosyncratic responses, different personal associations, cultural background and so on. It has been suggested that all these dimensions form the whole of a meaning which might therefore be distributed over multitude of brain areas. Semantic representations seem not to be confined to some meaning-specific brain regions but appear rather distributed in a systematic way throughout the entire brain (González et al., 2006, p. 909). This is in line with other studies as for instance Martin & Chao (2001) who suggest that information about different object features may be stored in different regions of the cortex. They found distributed representations in which the same neural systems that are active during perception are involved as well.

González et al. (2006) identify as one of the central issues in cognitive neurosciences unraveling the way in which words and their meanings are represented and processed in the brain. The research of meaning representations of other domains can surely contribute to this crucial point in that it can help to decipher its actual organisation in the brain and point out the overlapping representations with words, that are supposedly processed by distributed neural assemblies with cortical topographies that reflect their meaning or aspects of their reference.

Kölsch (2011) suggests that posterior temporal cortical regions might store conceptual features rather than only lexical representations per se which would explain why these regions can be activated by musical concepts. This effect might be accounted for by the mechanism of spreading activation. The chords, stimuli that were used in this study as a prime, may activate affective representations, which spread onto affectively related lexical representations (Kölsch, 2011, p. 93). It has also been pointed out that semiotic relations of musical events seem to be processed by the brain in a similar way as corresponding relations in the case of concepts coded by linguistic constructions, a fact indicating shared neural networks involved in common processing of language and music (Reich, 2011, p. 120).

Huth et al. (2006) have conducted an experiment showing intricate patterns of brain activity in response to information about specific semantic domains consistent across individuals. The domains represented groups of related concepts with areas selective for concrete or abstract, action words, social narratives etc. This associationist approach advances a holistic brain activation during comprehending disregarding the areas that are traditionally thought to be centres of specific linguistic processing. A cortical map of lexical features of words forms a vast semantic network in high-associative areas. No similar experiment has been conducted with musical material and the comparison of the results would probably be most revealing, although it is not easy to judge whether the vagueness of musical sense would permit replicable data.

It seems that musical meaning is at least partly processed with the same mechanisms as meaning in language and that the two domains are not strictly separate with respect to the processing of meaning (Kölsch, 2011, p. 102).

6. Evidence for Shared Processing

A lot of attention has been payed to analogies between syntax processing in the two domains. Since language and music draw on a common pool of limited resources it seems plausible that tasks combining linguistic and musical syntactic integration will show interference between the two. The pioneering works of Stefan Kölsch have indeed revealed these effects. First of all it was the ERP component ERAN (early right anterior negativity), musical equivalent of linguistic ELAN (early left anterior negativity) mirroring initial build-up of syntactic structure. ERAN is measurable rather on the right lateralised electrodes and is associated with processing of harmonic syntactic incongruities. The data show that ERAN might originate in Broca's area and its right hemisphere homologue. This finding suggests that musical syntactic processing is capable of activating some areas that are purportedly responsible for language processing. Kölsch et al. (2005) addressed the question what would happen when both kinds of incongruities – musical and linguistic – occur at the same time. In the critical condition, in which a sequence had simultaneous structural incongruities in language and music, an interaction was observed. The ELAN to syntactically incongruous words was significantly smaller when these words were accompanied by an out-of-key chord, the data suggesting therefore that processes underlying ELAN and ERAN are competing for similar neural resources. Most importantly, in a control experiment, it was revealed that this outcome was not due to general attentional effects because the size of the ELAN was not affected by a simple auditory anomaly manipulation that involved a physically deviant chord on the last word in a sentence. Similarly study of Schön & Morillon (2019) also showed that ELAN can be affected by melodic or harmonic unexpected events. The interactions provide an evidence for shared syntactic processing driven at least partly by shared neural resources (given the limited scope of this article details cannot be given, but for similar results see also Fedorenko et al. 2009).

As for the semantic processing, the data is rather sparse. The N400 ERP component that is thought to reflect the processing of extra-musical meaning has been found during the experiments of musical priming where the target words – concrete or abstract noun, followed semantically mismatched versus matched musical excerpts (Kölsch et al., 2004). Despite the fact that the specificity of semantic concepts activated by music is likely much lower and variable between listeners than those activated by language, the same effect as with linguistic prime has been obtained and both conditions activated the similar regions of the brain, namely the posterior portion of the middle temporal gyrus (MTG) and Brodmann areas (BA) 21 and 37. The amplitude and latency were for both linguistic and musical conditions equal. The existence of N400 for mismatched word following a musical prime indicates the capacity of music to convey a semantic reference or concept and to influence the semantic processing of words. Music is therefore in position to systematical activate semantic concepts and subsequently to convey a semantic information.

Daltrozzo & Schön (2008) have also found a larger N400 for word following a conceptually unrelated compared to related musical context. The same effect was found for musical targets, result that suggests the ability of music to convey concepts. It should be noted that concepts can be generated by other sources as for instance odours, yet it is not clear whether the storage and retrieval of such concepts are independent of the domain they are carried by. Since the N400 did not differ significantly between the two conditions the data argue more in favour of an a-specific cognitive module of conceptual processing or at least of a strong overlap of the cognitive modules of linguistic and musical concepts and hence an overlapping lexical access of linguistic and musical concepts. Attention has been drawn to the fact that musical lexicon stores representations of musical features quite like the word representations do. Hearing a novel melody can activate a composite pattern of musical features in the mental lexicon, and thus activate the corresponding representations, one of which might be at the conceptual level. These overlapping representations would thereby account for the overlapping lexical access of linguistic and musical concepts (Daltrozzo & Schön, 2008, p. 1890).

Kölsch et al. (2004) further advances that it is possible that semantic concepts as mental representations are stored without having the recourse to language, and are therefore activated by music as well as by words. Accordingly the musical information would activate conceptual representations per se rather than its covert verbalisation (Kölsch, 2011, p. 92). It has therefore been suggested that musical and language semantics are processed in partially overlapping regions and neural resources, through the same cognitive processes. At the same time Kutas et al. (2010) underscore the fact that the N400 data point to a distributed, multimodal and bihemispheric comprehension system that is simultaneously open to linguistic and nonlinguistic influences, which often interact. This suggests again a predictive, flexible and context-dependent comprehending.

By the same token Kölsch et al. (2004) suggest that the overlapping neuronal network of music and language explains the important role the musical elements of language play early in the process of language acquisition and suggests that the musical elements of language precede the phonetic elements in the acquisition of linguistic competence. It had also been pointed out that speech is initially processed by infants as a type of music (Hodges & Thaut, 2019, p. 12).

That being said, one should bear in mind that neural overlap does not necessarily entail neural sharing since distinct brain pattern of responses may be found within overlapping regions. It could also be that different patterns of the same neural population account for their different reaction to music and language. Most importantly similarity of activation does not imply that the underlying processes are similar, or that the recruited areas are critical to performing a given task (Peretz & Zatorre, 2005, p. 91).

7. Meaning on the Verge of Syntax and Semantics

The psychological reality of music-semantic processing is reflected in ERP components. Several different classes of meaning (see Kölsch et al., 2004) are mediated by different cognitive processes and reflected in different ERPs. Besides the N400, another important component has been found that is thought to reflect intra-musical meaning that emerges from the cognitive interpretation of intrinsic relations between musical sounds in a sequence, the N5 (Kölsch, 2011, p. 89). It is supposed to reflect general principles of meaning emerging from structural relations which are the same as for poetry, visual arts etc. This finding supports the idea that intra-musical phenomena can give rise to extra-musical meaning since meaning is emerging from harmonic integration due to the construction of structural model. As a matter of fact it seems that semantic aspects of music arise from the formal ones and therefore depend on the structuring of the material (Dürr, 2004, p. 255). Indeed music seems meaningful because it is rich in perceptual information and because it is organised and its progress is predictable (Davies, 2011, p. 114). One of the important features of musical meaning is the fact that listeners acquire over time sensitivity to the statistical distributions of tones within the music and can infer some structural relations on this basis. As a consequence expectation becomes an important process in creating formal meaning procuring the sense of logical connectedness, progress and direction.

By the same token in language processing the N400 reflects the process of semantic integration of the critical word with the working context and results therefore from combinatorial process, not simple lexical-level processes (Lau et al., 2008, p. 921). It has also been suggested that ERPs reflect activation of a general-purpose violation detector that operates in a variety of domains which means that similar cognitive operations may be involved in processing the structure of language and other types of well-organised, rule-based knowledge, such as music (Besson et al., 1998, p. 494).

Summarising the results of their priming experiment, Steinbeis & Kölsch (2008) put forward that the found interaction between harmonic structure processing and semantic linguistic material suggests that musical structure is a key feature leading to semantic processing and making sense of music. In the same vein the structure of a normal linguistic utterance can also give rise to certain semantic expectations, for instance the membership of a certain word class. Steinbeis & Kölsch (2008) evoke therefore the theoretical possibility that structural violations may be of both a syntactic and a semantic nature.

Moreover Wassiliwizky & Menninghaus (2021) advance that an important feature of meaning in poetic language is the parallelistic patterning, e.g. repetitive structures. Should they be removed while the semantic content remains, the emotional involvement and overall enjoyable affective response would be reduced. The structure plays, here again, an important role since the higher-order properties of a stimulus result from an internal integration of basic features into more complex semantic concepts which do not have to be consciously accessible to the observer.

Given the diversity of tasks eliciting N400 (see e.g. Pylkkänen et al., 2009; Patterson et al., 2007; Friederici, 1993; Friederici, 2000) a suggestion has been made that largely overlapping widespread networks are involved in processing syntactic and conceptual anomalies. One could argue that these findings might reflect some kind of interference between the two processing streams thus suggesting a convergent feature with processing of musical meaning, which relies for the major part on the syntactic and logical structural sequences. Kölsch & Schröger (2007) put forward that both music and language meaning require the establishing of relationships between successive events in order to process the structural aspects of sequential information, as well as detection of structural irregularities. Creation of syntactic structure – a mental representation of how the elements are organised to form a hierarchically nested structure – may then produce a musical semantic structure – a mental representation of how the meanings of individual events are combined to yield the meaning of entire sequence (definition of syntactic and semantic structure adapted for music from Lau et al., 2008). Rather than purely semantic anomaly, the N400 might reflect violations of rules that are relevant for interpretation. Since N400 appears for violations of concepts as well as for violations of grammatical rules (contextual violations, compositional semantics violations, grammatical violations, implausible continuations) it is plausible that semantics is shared between grammar and extra-linguistic representations, and the semantic processing relies on distributed network with different hubs, and problem occurring at any of these elicits the N400 effect. Therefore, the two domains would again rely on the same processing resources making the interferences possible.

Fitch & Gingras (2011) make a good point by stating that there is no need to draw a clear line between 'semantics' and purely structural 'syntax' in music. By the same token Daniel Everett[2]

2 Personal communication

advances that language should be treated as a system of complex interactions between syntax, semantics and pragmatics. If we consider language from the evolutionary point of view, and agree with the hypothesis that language has evolved for the communication purpose, then the interconnectedness of all its parts seems obvious.

8. Effects of Music on the Listener

Pehrs et al. (2014) investigated the way human brain integrates the visual information and the musical soundtrack into a coherent percept while watching movies and found out that evaluation of the same scene can be altered by emotionally loaded music. The emotions induced by music can therefore have a modulatory impact. Although song is not a multisensory stimulus, it is nevertheless a composed stimulus, a fact allowing the assumption, that a similar modulation might be in act in song listening. The valence of accompaniment might for instance modulate definition of polysemous words.

In fact music elicits euphoria and craving similar to tangible rewards that involve the striatal dopaminergic system as well as the mesolimbic system implicated in reinforcement and motivation in response to biological reinforcers necessary for survival (Salimpoor et al., 2011, p. 257). For this reason people experience an intense pleasure in response to music due to the dopamine release. These effects on arousal and mood constitute an important component of the meaning of music (Fitch, 2006, p. 179). In fact there is evidence of broad agreement among listeners in judging expressive qualities of music. Kölsch et al. (2005) develop by pointing out that any emotional activity is always related to bodily reactions whose perception or awareness conveys a sense of meaning. It seems highly plausible that such a strong physiological reaction is likely to impact on the simultaneously presented linguistic stimulus. What's more Taruffi et al. (2017) suggest that sad music engages the listener in mind-wandering and is able to disrupt ongoing task performance by disengaging of attention from perception.

9. Evidence for Independent Processing

Let us briefly summarise evidence against the view of music and language as integrated into a single percept in the context of a song. Besson et al. (1998) argue in favour of independent modules of processing of the two domains. In their study, when an incongruous word (with regard to the preceding context and reflected by the N400) was sung on an out-of-key tone (reflected here by the P300) the effect was additive which suggests that semantic and harmonic violations are processed independently. Nonetheless it could be argued that a harmonic violation (as an out-of-key chord) elicits a component sensitive to the surprise and attention shift rather than a semantic incongruity. As a matter of fact the P300 is known to pop up in experiments using the oddball design which requires the reorientation in the response to the stimulus that attracts the attention.

Furthermore the modular organisation of the human cognitive system which constitutes the theoretical framework of this study has proven problematic and has been debated in the recent years. It seems rather that no complex cognitive process can be performed by a single brain area. Although we might possess convincing evidence for a specific role that a certain region plays, it does not mean by far that other regions are not equally involved. Indeed Cobb (2021) insist that the brain does not seem to be organised into independent modules in the same way as a computer is. Neurones and the networks they form are highly interconnected and affect the electric activity and genetic expressions of neighbouring regions.

Grounding on the assumption that listening to songs may represent a typical case of divided attention, Bonnel et al. (2001) used the dual-task paradigm in order to find out whether attention to one source of information adversely affects the perception of the other and to assess the potential patterns of interference. The results suggest that the participants were able to divide their attention between lyrics and tunes which demonstrates that the processes involved in performing the semantic and melodic detection tasks did not compete for the same pool of resources (Bonnel et al., 2001, p. 1210). That being so, perception of one component of a song seems unaffected by perception of the other. This argues against the assumption that the semantic and melodic components were integrated within a unified percept and are therefore processed by independent systems. In the same vein Peretz & Zatorre (2005) suggest the existence of neural networks specialised for the processing of scale structure in melodies, a fact that would at least partially explain the results.

The researchers reason therefore that a song might not be perceived as a single object having two dimensions, but rather composed of two separate objects that have one dimension each. This conclusion is nevertheless at variance with the study of Zatorre and Samson (1999), where it was found that when remembered, the song is stored with the words, since without them the recollection is much harder. At the same time Poulin-Charronat et al. (2005) have shown that unexpected harmonies increase response times on a lexical decision task in a semantic priming paradigm.

Furthermore, what has actually been tested was interference of a contextually incongruous word with a deviant tone of a melodic contour, not musical semantics per se. The only musical aspect was here the melodic contour which might not have been sufficiently rich enough to convey a concept. Hence, the results might reflect a reaction to an unexpected event and the attentional shift rather than semantic processing in itself.

On the other hand it has to be pointed out that neuropsychological investigations of patients with selective brain damage have revealed cases of double dissociations between language and music. Consequently it seems that at least some aspects of music and language processing are indisputably implemented in different areas in the brain. Current evidence suggests the existence of specific neural networks devoted to some aspects of musical and language processing, as corroborated by the clinical evidence from patients who preserve their musical abilities while exhibiting profound disturbances in language, as well as the opposite pattern in patient affected by amusia while linguistic processing is spared (Peretz et al., 1994). Evidence from non-fluent aphasics impaired at recognition of spoken words while spared recognition of music also suggests the functional independence of music and speech at some levels (Peretz & Zatorre, 2005, p. 106).

10. Overlapping and Independent Functions

One of the Darwin's core hypotheses has recently been brought to light, namely the theory of musical protolanguage (Fitch, 2010, p. 50). An a-referential songlike communication system might have predated language in human evolutionary history. This precursor of human language was not a vehicle for the unlimited expression of thought, but rather an emotion-based communication system. Modern music might therefore exist as a sort of behavioural fossil of this past system. Because the neural mechanisms underlying song were precursors of phonological mechanisms in spoken language, considerable overlap is expected between phonological and musical abilities. On the other hand language-specific mechanisms especially those involved in semantics are probably independent.

At the same time, there is a copious evidence for innate neural structures supporting music (Hodges, 2019, p. 25). Indeed McDermott & Hauser (2005) suggest that music perception is constrained by innate, possibly human- and music- specific principles of organisation, and for these putative music-specific neuronal networks speak other studies as well (eg. Peretz & Zatorre, 2005). Although a full overlap of semantic processing is unlikely since music conveys different kinds of concepts than language which do not necessarily require verbalisation, it has been suggested that music, despite being an autonomous innately constrained function, is made up of multiple modules that overlap with other functions such as language (Peretz, 2006, p. 25).

In the study of Sammler et al. (2010) the left mid-STS (superior temporal sulcus) showed an interaction of the adaptation effects for lyrics and tunes, suggesting an integrated processing of the two components at prelexical, phonemic processing level. Lyrics and tunes seem to be processed at varying degrees of integration and separation through the consecutive processing levels following a gradient from more to less integrated processing along the posterior-anterior axis of the left STS and the left PrCG (cingulate gyrus) (Sammler et al., 2010, p. 3576). Beyond that, the data suggest a separate processing of lyrics at subsequent levels of structural analysis and lexical-semantic representations in the anterior part of STS which suggests a greater autonomy of linguistic meaning in songs. The lack of interaction and the stronger adaptation for lyrics than for tunes has been interpreted as an independent processing of lyrics, perhaps resulting from the processing of meaning. The results indicate that meaning represents precisely the point where the two domains diverge in different directions.

The question of overlapping mechanisms is fairly intricate and Schön & Morillon (2019) discourage from thinking about language and music as a whole, but rather in terms of precisely defined elementary operations. Indeed, some of these might be overlapping and some might be subserved by specific circuits. Music is not a unitary cognitive module since different components of 'the music faculty' may have different evolutionary histories (Fitch, 2006, p. 174). In order to be fully examined, every complex behaviour needs to be broken into independent components. A promising approach seems to study whether two corresponding levels of music and language processing interact or not. Most importantly it appears that the degree of integration or separation of the two domains depends on the specific cognitive processes targeted by an experimental task (Sammler et al., 2010, p. 3577). Undoubtedly one of the biggest challenges for the research on the music-language interface is to draw further analogies at higher levels and to find comparable units. Music and language need to be reduced to more elementary functions and it seems that the right semantic primitives of both have not yet been found. It will take a great deal of further research in order to give a satisfactory answer.

That being said it should be emphasised that the results depend crucially on the task. For instance Samson & Zatorre (1999) admit that the two components of a song could be encoded as a single event in memory for song but have found evidence for the use of dual memory codes suggesting a different role of each temporal lobe in memorising songs. In the case of a song, the words might actually form part of the stimulus that defines the tune. The findings suggest that the verbal and melodic codes are partially integrated, but words can be to a certain degree independently accessible.

By the same token it is sometimes difficult to determine with precision the nature of overlaps. Fedorenko et al. (2009) propose another explanation for the seemingly competing and overlapping structure building processes of music and language. It is possible that there is rather an overlap at the retrieval stage of language processing, and that 'at the stage of retrieving the linguistic representation from memory, the presence of a complex structural integration in the accompanying musical stimulus makes the process of reconstructing the syntactic dependency structure more difficult' (Fedorenko et al. 2009, p. 6).

11. Conclusion

On one hand the musical meaning seems to be understood in terms of physiological changes that listening triggers by inducing emotions as well as with reference to engendered impressions of movement through space. On the other hand it seems to rely on learned probabilistic statistics of pattern sequences which create expectations that are subsequently experienced as meaningful. It would therefore be the sense of logical structure, maybe comparable to linguistic syntax, that bestows on music the referential power. It has been shown that music is not semantically deficient relative to language, yet it encourages a complementary mode of interpretation which might be precisely the major source of its appeal. An array of experimental studies has been brought together showing clearly the ever growing interest of researchers coming from diverse domains who all see the research on music-language interface as a unique opportunity to better understand the human cognitive capacities and organisation of the human brain. It has been shown that the processing of musical and linguistic meaning probably rely partially on shared and partially on specific mechanisms. It has been demonstrated that studies with patients with selective brain damage point rather in the direction of separate processing modules whereas studies using ERPs have yielded results inconsistent with domain-specific view. The precise extent of domain specificity of language and music still remaining elusive. Nevertheless, since both domains are uniquely human and therefore cannot be examined in animals, (as for instance the memory and emotions) a joint research can strongly contribute to the knowledge of each separate domain. In order to fully grasp their organisation there is a great need for an interdisciplinary approach. Recently, one of the factors prompting research in the domain of music and language is the possibility to use music to remediate language impairment, e.g. to recover impaired functions by defining what aspects of music training benefit language processing and at which levels of processing.

12. References

Besson, M., Faïta, F., Peretz, I., Bonnel, A.-M., Requin, J. (1998). Singing in the Brain: Independence of Lyrics and Tunes. *Psychological Science*, Vol. 9, No. 2.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences*, *930*, 232–258. https://doi.org/10.1111/j.1749-6632.2001.tb05736.x

Bonnel, A-M., Faita, F., Peretz, I., Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics 63* (7).

Brogniez, L. & Piret, P. (2005). Introduction. In *Textyles* 26-27. Retrieved from http://textyles.revues.org/462

Cobb, M. (2021). Une brève histoire du cerveau. Dunod.

Cross, I. (2011). The meanings of musical meanings. Comment. *Physics of Life Reviews 8*. 116- 119

Culpeper, J. & Semino, E. (2002). Cognitive Stylistics : Language and Cognition in Text Analysis. John Benjamins Publishing Co. Amsterdam.

Cusimano, C. (2015). Le Sens en mouvement - Études de sémantique interprétative. 1st edition. Frankfurt am Main: Peter Lang, 2015. 137 s. Etudes de linguistique, littérature et art. Vol. 11. ISBN 978-3-631-66014-0.

Daltrozzo, J. & Schön, D. (2008). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of cognitive neuroscience 21:10*, pp. 1882-1892. MIT

Davies, S. (2011). Questioning the distinction between intra- and extra-musical meaning. Comment. *Physics of Life Reviews 8*. 114-115

Dürr, W. (2004). Sprache und Musik. Bärenreiter, Kassel.

Fedorenko, E., Patel, A., Casasanto, D. *et al.* (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition* **37,** 1–9. https://doi.org/10.3758/MC.37.1.1

Feldman, J. & Narayanan, S. (2004). Embodied meaning in a neural theory of language. *Brain and Language 89*. 385-392

Fitch, W. (2006). The Biology and Evolution of Music: A Comparative Perspective. *Cognition*, vol. 100, Elsevier. 173-215.

Fitch, W. (2010). Musical protolanguage. In *The Evolution of Language* (Approaches to the Evolution of Language, pp. 466-507). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511817779.015

Fitch, W. & Gingras, B. (2011). Multiple varieties of musical meaning. Comment. *Physics of Life Reviews 8*. 108-109

González, J., Barros-Loscertales, A., Pulvermüller, F. (2006). Reading *Cinnamon* Activates Olfactory Brain Regions. *Neuroimage* 32. 906 – 912

Grandjean, D. (2021). Module 4 – Perception de la prosodie linguistique et émotionnelle [MOOC lecture]. In D. Grandjean, N. Golestani,, D. Marie, *Neurosciences : audition, langage et musique*. https://cours.edulib.org/courses/course-v1:UMontreal+BRAMS.4+H2020/course/

Gribenski, M. (2005). Musique et sciences cognitives. In *Labyrinthe*, 20 | 2005 (1). Retrieved from http://journals.openedition.org/labyrinthe/763 ; DOI : 10.4000/labyrinthe.763

Harmat, A-M. (2010). Formes musicales et formes littéraires. Réflexions sur le dispositif esthétique musical au service de la représentation littéraire. In *Fabula / Les colloques, Littérature et*

musique. Retrieved from http://www.fabula.org/colloques/document1267.php
Hodges, D. (2019). Music Through the Lens of Cultural Neuroscience. In M. Thaut & D. Hodges (Eds.), *The Oxford Handbook of Music and the Brain*. Oxford University Press.
Hubbard, T.L. (2019). Neural Mechanisms of Music Imagery. In M. Thaut & D. Hodges (Eds.), *The Oxford Handbook of Music and the Brain*. Oxford University Press.
Huth, A., de Heer, W. et al. (2006). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* vol. 532. 453-472. doi:10.1038/nature17637
Jentschke, S., Kölsch, S. & Friederici, A. (2005). Investigating the Relationship of Music and Language in Children. *Annals New York Academy of Sciences 1060*: 231–242. doi: 10.1196/annals.1360.016
Kölsch, S et al. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature Neuroscience 7*(3)
Kölsch**,** S. (2005). Ein neurokognitives Modell der Musikperzeption. *Musiktherapeutische Umschau 26*,4. ISSN 0172-5505
Kölsch S. (2007). Musik verstehen – eine neurowissenschaftliche Perspektive. In: Becker & Vogel: *Musikalischer Sinn*, Suhrkamp.
Kölsch**,** S. & Schröger E. (2007). Neurowissenschaftliche Grundlagen der Musikverarbeitung. Retrieved from
Kölsch, S. (2011). Towards a neural basis of processing musical semantics. *Physics of Life Reviews 8*. 89-105
https://stefan-koelsch.de/papers/Koelsch-Schroeger-Musikpsychologie_20070221.pdf
Kutas, M. & Federmeier, K. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-related Brain Potential (ERP). *Annu. Rev. Psychol.* Retrieved from www.annualreviews.org
Lau, E., Phillips, C., Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature*, vol. 9, 920-933. doi:10.1038/nrn2532
Lerdahl, F. (2003). The sound of poetry viewed as music. 2003. In I. Peretz & Zatorre, R. (Eds) *The Cognitive Neuroscience of Music*. Oxford University Press.
Limb C. J. (2006). Structural and functional neural correlates of music perception. *The anatomical record. Part A, Discoveries in molecular, cellular, and evolutionary biology*, *288*(4), 435–446. https://doi.org/10.1002/ar.a.20316
Locatelli, A. & Delpy, J.-A. (2009). Littérature et musique : points d'achoppement et de rencontre. *Québec français, Littérature et musique* (152), 30–33. Retrieved from https://id.erudit.org/iderudit/44182ac
Martin, A. & Chao, L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, 11: 194-201.
Martinec, R. (2018). Linguistic Rhythm and its Meaning: Rhythm Waves and Semantic Fields. *Linguistics and the Human Sciences*, Vol. 14(1/2), p.70
McDermott, J. & Hauser, M. (2005). The Origins of Music: Innateness, Uniqueness, and Evolution. *Music Perception*, Vol. 23 September
Molino, J. (2009). Le Singe musicien. Actes Sud / INA.
Nieuwland, M. & Van Berkum, J. (2006). When Peanuts Fall in Love: N400 Evidence for the Power of Discourse. *Journal of Cognitive Neuroscience* 18:7, pp. 1098-1111. Massachuesetts Institute of Technology
Nöth, W. (1995). Handbook of semiotics. Indiana University Press.
Patel, A. (2008). Music, Language and the Brain. Oxford University Press.
Pehrs, C. et al. (2014). How music alters a kiss: STG controls fusiform-amygdalar effective connectivity. SCAN 9, 1770-1778. doi:10.1093/scan/nst169
Patterson, K., Nestor, P., Rogers, T. (2007). Where do we know what we know? The representation of semantic knowledge in the human brain. *Nature*, vol. 8. 976-988. doi:10.1038/nrn2277
Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology 10(1)*:21-56 DOI: 10.1080/02643299308253455

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain : a journal of neurology*, *117 (Pt 6)*, 1283–1301. https://doi.org/10.1093/brain/117.6.1283

Peretz, I. (2006). The Nature of Music from a Biological Perspective. *Cognition* 100, 1-32. doi:10.1016/j.cognition.2005.11.004

Peretz, I. & Zatorre, R. (2005). Brain Organization of Music processing. A*nnual Review of Psychology*, 56(1):89-114. DOI: 10.1146/annurev.psych.56.091103.070225

Poulin-Charronnat, B., Bigand, E., Madurell, F., Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition* 94(3):B67-78. doi:10.1016/j.cognition.2004.05.003

Pylkkänen, L., Oliveri, B., Smart, A. (2009) Semantics vs. world knowledge in prefrontal cortex. In *Language and Cognitive Processes* 1-22. Psychology Press. 20 DOI: 10.1080/01690960903120176

Rasse, C., Onysko, A., Citron, F. (2020). Conceptual Metaphors in Poetry Interpretation: A Psycholinguistic Approach. *Language and Cognition*. DOI: 10.1017/langcog.2019.47

Reich, U. (2011). The meanings of semantics. Comment. *Physics of Life Reviews 8*. 120-121

Salimpoor, V., Benovoy, M., Zatorre, R. et al. (2011). Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music. *Nature Neuroscience* 14(2):257-62. DOI: 10.1038/nn.2726

Sammler, D., Baird, A., Valabrègue, R., Clément, S. et al. (2010). The Relationship of Lyrics and Tunes in the Processing of Unfamiliar Songs: fMRI Adaptation Study. T*he Journal of Neuroscience* 30(10):3572–3578. DOI:10.1523/JNEUROSCI.2751-09.2010

Samson, S., & Zatorre, R. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: evidence for dual encoding. *Journal of experimental psychology. Learning, memory, and cognition, 17 4*, 793-804 .

Schön, D. & Morillon, B. (2019). Music and Language. In M. Thaut & D. Hodges (Eds.), *The Oxford Handbook of Music and the Brain*. Oxford University Press.

Sebeok, T. (2001). *Signs, An Introduction to Semiotics*. University of Toronto Press.

Steinbeis, N., & Kölsch, S. (2008). Shared Neural Resources between Music and Language Indicate Semantic Processing of Musical Tension-Resolution Patterns, *Cerebral Cortex*, *18(5)* 1169–1178

Taruffi, L., Pehrs, C., Skouras, S., Koelsch, S. (2017). Effects of Sad and Happy Music on Mind-Wandering and the Default Mode Network. *Scientific Reports, 7*: 14396. nature.com/scientificreports

Vallespir, M. (2010). Langage et musique : approches sémiotiques, In *Fabula / Les colloques, Littérature et musique.* Retrieved from http://www.fabula.org/colloques/document1274.php,

Wassiliwizky, E. & Menninghaus, W. (2021). Why and How Should Cognitive Sciences Care about Aesthetics? *Trends in Cognitive Sciences 25(6)* DOI:10.1016/j.tics.2021.03.008

Correspondence details

Alice Karbanova
Masaryk University, Brno, Czech Republic
alice.karbanova@gmail.com