

Linköping University Postprint

The Words of Music

Johan Fornäs

N.B.: When citing this work, cite the original article.

Original publication:

Johan Fornäs, The Words of Music, 2003, Popular Music and Society, (26), 1, 37-51.
<http://dx.doi.org/10.1080/0300776032000076388>.

Copyright © [Taylor & Francis Group](#), an informa business

Postprint available free at:

Linköping University E-Press: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-15568>

The Words of Music

By Johan Fornäs

Professor at the National Institute for Working Life:
Work & Culture Program in Norrköping, Sweden

Mail address (home): Fatbursgatan 18A, S-118 54 Stockholm, Sweden

E-mail: johan.fornas@niwl.se

Phone: +46-8-6444308 or +46-11-218904

Johan Fornäs has a background in musicology and in media & communication studies. He is now Professor of Cultural Production Studies at the Work & Culture Program of the National Institute for Working Life in Norrköping, Sweden. His many publications about rock, alternative movements, youth culture, identity, modernity, media use and cultural theory include the three 1995 English volumes *In Garageland: Rock, Youth and Modernity*, *Youth Culture in Late Modernity* and *Cultural Theory and Late Modernity*.

The Words of Music

Words and music are two distinct symbolic modes. Yet, as human ways of communication or forms of expression, they have much in common. They have important similarities as signifying systems, their mutual penetration is suggested by metaphors like “the language of music” or “the music of language”, and they are multi-modally united in all song genres. Popular song in particular is generally “text-intensive” (Booth 188). Many attempts to distinguish univocally between them turn out to be more difficult than may first be expected. Most cultural practices have no problems separating or combining them, but their defining differences tend to be explained in highly divergent and even contradictory ways. I will here exemplify some such fascinating paradoxes, problematizing what might seem to be self-evident rather than offering any simple answers. By pointing at certain interesting complexities, at least some main dimensions of the relations between words and music will be discerned.¹

How is the difference between music and words constructed in cultural practices and theories? The first section mentions some obscure aspects of the words/music distinction in listening practice, with an emphasis on popular song. The general juxtaposition of the verbal and the musical symbolic modes, as two institutionalized fields of cultural practice and theorization, is particularly dramatized in the phenomenon of song, where their coexistence in one single performance act puts their distinction to a difficult test. Digitally processed techno and rap music has sometimes deliberately blurred the boundaries between speech and song, between words and sounds or between lyrics and music. The second section presents a few such confusing musical examples that experiment with those borders, by transforming words and noises into music, and making traditional generic classifications useless. The third section then discusses some general problems in conventional definitions of the basic terms, while the final section sums up a possible way to regard the bifurcation of the verbal and musical modes.

Words with music

Song is a multimodal supergenre that mediates between words and music.² In most rock and pop songs, they are treated as neatly distinct. On record sleeves,

the formula “Words & Music” points towards the *creators* of a song, indicating that they may be two different persons or at least distinct functions: one author of the lyrics, and one composer of the melody and its accompaniment. In rock, these functions are often amalgamated, when a single singer-songwriter, a pair of artists or a whole band is presented as the undifferentiated origin of both words and tunes. But copyright laws and music industry practices still tend to stick to this dichotomy, and even singer-songwriters or tight rock groups sometimes adhere to it by writing “Words & Music: ...” rather than just “Songs made by: ...”.

However, the “writing” of any song-hook, like for instance “Be bop-a-lula”, often means inventing words and music simultaneously, in one single move. Such a text-line when spoken is as much a rudimentary melodic motive as are the appropriate five notes on a musical staff. It creates a rhythmic organization that makes only a certain range of musical realization possible. And in the actual performance of such a hook (or indeed any song), the singer again performs the words and the music as one unified whole, in one single act which is indivisible in time and physical space. Instruments play the music, but the singer performs lyrics and tunes absolutely simultaneously. So, the distinction between writing (or performing) music and writing (or performing) words is not always quite clear.

At the opposite end of the communicative process, *listeners* also often tend to differentiate between words and music, even though they are reached the ear at the same moment. Being socialized into modern listening practices, most people can instantaneously disentangle the lyrics from the melodic and rhythmic lines, interpret the words and compare that meaning with the musical sounds by which it is supported (or counteracted). One may disregard or actively listen for the words, feel how they combine with the music, but it is hard not to hear them as something other and more than pure musical sound. Discussions of popular songs generally make those distinctions: “Good tune, but lousy lyrics”....

But again, things are not actually quite so clearly divided. Listening to songs and singing means entering a mode of perception where words and music continually interfere. To understand a verbal text one has to perceive how its units are articulated and grouped, which is affected by its rhythmic and melodic performance. And the sound, rhythmic and melodic parameters of the music are in turn depending on which words that are sung. The material qualities, the form-relations and the semantics of words and music often tend to merge. The

verbal content and the musical organization are thus linked by an intense cross-traffic, rather than being the two completely separate systems they often appear to be in the end – and in much theoretical analysis.

The different ways to *translate* aural into visual forms further emphasize their bifurcation. It mostly is possible to transcribe and analyze the main musical structure of a pop tune approximately as some kind of a score, with the words in alphabetic writing under the melody. But even though such a notation system creates a neat division, it is never the same as the actual song. Every such visual and spatializing translation of an aural and temporal performance is a kind of analysis that separates what was united by using fundamentally different kinds of transcription systems for the words and the music. In modern Western societies, both the alphabet and conventional musical notation utilize reductively discrete means to summarize complex and continuous processes. They both translate time flows into vertically stacked (read downwards) horizontal lines (read from left to right), written or printed on pages (read in a routine order from front to back). But they still obviously differ in the precise way this is done, and when they are combined, it is generally easy to see which visual signs are to be read as verbal and which belong to the musical level.

Other forms of fixation – as for example the graphs produced by an oscillograph or the engraved tracks for the needle or the laser beam in a vinyl or a CD record – do not clearly separate words from music within the continuous sound flow of a song. Several elements of a song seem to be simultaneously words and music; in fact, this is true of most of them! Even simple common words have duration and pitch whether sung or spoken, and within the frame of a song these can be interpreted as musical parameters. Conversely, the most nonsensical utterance can hypothetically be understood and transcribed as part of its lyrics. The separation between words and music is thus only the result of a complicated but quasi-automatic analytic process effected by the transcriber already in the first listening.

Song thus contains both words and music, but speech performance is also more than just a neutral deliverance of verbal semantics. In a rich key work on music and performance, Simon Frith (1996: 159) argues that three things are heard at once when listening to the lyrics of pop songs: *words* (as a “source of semantic meaning”), *rhetoric* (“words being used in a special, musical way”) and *voices* (“human tones” as “signs of persons and personality”). Performance traits like vocal gestures, timbre and rhythmical inflections are always and

inevitably present in spoken words, but they tend to be perceived as some kind of addition to the words as such, and to disappear when they are transcribed in writing. Like song adds a melody to song lyrics, speech seems to superimpose certain essentially non-verbal traits to the words spoken. Speech and song are two modes of vocal performance, and while singing written words transforms texts from writing to song, reading them loud reconstructs them as speech. Visual markers and signs are in both cases translated into audible sound structures. Combinations of letters are read as combinations of phonemes, while intonation and phrasing helps conveying the formal organization represented by dots and commas in the written version. Reading written words loud is thus, just like singing, a performance mediating between the aural and the visual symbolic modes. “The voice records a text”, and “a song text is a script for a public event” (Booth 187 and 34), but this is as true when a text is performed in speech form, in lectures, radio shows, live poetry or theatre drama.

One can actually choose to listen to ordinary speech *as if* it was a song! Instead of interpreting the rhythmic, melodic and timbre aspects of a voice and a speech performance in linguistic terms, they can be decoded as music. It would be more or less difficult, depending on the type of speech, and it is not a form of reception that is encouraged and preferred by the speaker or the contextual setting. But it is always possible, since speech uses the same oral medium as song, including sounding words that can be analyzed in terms of many of the same parameters that are used to understand musical works. Many song genres include “heightened” speech – like opera recitative, Brechtian speech-song, speech choirs, raï or rap – and they are usually classified as speech-like song well integrated into a musical web rather than as non-musical speech that just happens to have a musical background.

Machine noises can be sampled and re-performed in a concert hall as musical sounds. This was done by art music futurists in the early 20th century and by avantgarde rock/pop groups half a century later. A steelworker with a musical ear may also perceive the surrounding factory sounds as having a musical quality. A poetic person might similarly perceive speech or even non-human sounds like street traffic, bird song, waves or the wind in the trees as having a musical quality.³ A performer can use various markers to steer the listener into decoding any sounds either as noises or as music, as speech or as song. In soundscapes (Schafer), film music (Gorbman 56ff), and muzak (Lanza

2f), the perception of sound sometimes balances or moves up and down across the border between noise, speech and music.

Meaning is made through acts of interpretation, in encounters between texts and subjects in specific contexts. It is through hearing sounds, rather than making them, that their meanings are produced, including the distinctions between genres and types of expression. Reception is here more essential than material production. Listening practices are no simple reproduction of encoded meanings, but a highly productive form of consumption, producing impressions, emotions, social relations and meaning. As the most distinctive way to use cultural phenomena or consume cultural commodities, interpretation is an active creation of meaning, resulting from the contextualized encounter between human subjects and texts. Independent of how they are produced, certain sounds are perceived in an “aesthetic” way, activating particular codes of interpretation and criteria of judgement which are situated in historically developed frameworks, including publishers, copyright legislators and music critics who uphold the standard division of words/music. John Blacking (10) puts a similar “emphasis on the primacy of listening”, as does Nicholas Cook (10): “Music is an interaction between sound and listener”.⁴ Having grown up in a culture where song and speech have for centuries been cultivated and institutionalized as separate symbolic modes, people are only so used to make instantaneous classification of sounds, that such automated distinctions are easily (though falsely) thought of as absolutely given.

Words as music

Everyday musical and analytical practices in modern Western society mostly take some kind of distinction between words and music for granted. However, some earlier, marginal or experimental kinds of music-making and listening either escape these entities or actively rework and stretch them. This is also true for recent popular music practices, in film scores as well as in rap and dance genres. Such limit cases point at the blurred character of the borderline between words and music, as well as between song and speech.

In 1993, the Swedish dance group L.P.C. (Lucky People Center) released a fascinating album called *Welcome to Lucky People Center*, with an intriguing opening. First, the sound of striking a match is followed by a blowing sound, as from a burning flame. Over this sound, a dense monologue is monotonously

spoken by the male singer's voice. The voice belongs to Freddie Wadling, who is the main singer of the L.P.C.. Half way through, another voice is also heard, slowly repeating "tic-toc" louder and louder until the last "toc" coincides with the word "bomb", after which this first number of the record ends with a sudden silence, just before the second "tune" starts. This first piece of just forty seconds is transcribed in the CD-booklet in a similarly condensed and uninterrupted way as it is performed:

he used to ask his mother why are we here where do we go how is it that you are dressed like that why am I not as big as you where is the sun when the moon shines why don't I breathe under water is the sum of a mass not equivalent to a form is santa claus a transsexual does rock'n'roll make you rich as the boy grew into manhood he kept asking all kinds of people about all kinds of things until his mother seriously considered retroactive abortion or to take the boy to a psychiatrist or even a taxidermist or to the dogfoodfactory but instead she bought the boy a microphone! tictoc poor old mum you are sitting on a bomb!!

This is the way "Tictoc" is transcribed on the CD cover. Another way of presenting its text in written form would be as follows.

He used to ask his mother:

- Why are we here?
- Where do we go?
- How is it that you are dressed like that?
- Why am I not as big as you?
- Where is the sun when the moon shines?
- Why don't I breathe under water?
- Is the sum of a mass not equivalent to a form?
- Is Santa Claus a transsexual?
- Does rock'n'roll make you rich?

As the boy grew into manhood
he kept asking all kinds of people about all kinds of things
until his mother seriously considered retroactive abortion
or to take the boy to a psychiatrist
or even a taxidermist
or to the dogfoodfactory.

But instead she bought the boy a microphone!

Tictoc...

Poor old mum – you are sitting on a bomb!!

While the album cover version expresses the seemingly unstoppable and chaotic flow of words as sounding materiality and deliberately leaves much ordering and interpretation work to the reader, this alternative rendering offers a more definite and easily discernable form-relational (syntactic) structure and (semantic) meaning of the text. One may perhaps say that the actual vocal

performance lies somewhere half way between these two versions. It is interesting to compare what the typography does to the written text with what the phrasing and emphases do to the spoken or sung text. Pictorial aspects of the script obviously interrelate with its meaning and impact, just like the “musical” aspects of the speech or song performance affect the way its words are interpreted by a listener. In both versions, materiality, form-relations and meaning interact densely, rather than being totally separated from each other. Meaning is affected by, yes even constructed only on the basis of material structures, while a reader’s or listener’s preliminary understanding of what the lyrics means in turn aids the perception of some of its material and formal aspects as well. The difference in how these two verbal transcriptions are experienced shows that modes of transcription are useful for analysis and understanding, but never innocently neutral. The same applies for musical transcriptions, as well as for genre classifications.

“Tictoc” may be perceived as a spoken introduction to the whole record album, or as a separate tune of its own. Its presentation on the record does not distinguish it from the other tunes, except for its shorter length. As a separate piece, it has traits reminding of theatre, like a radio drama. It also has similarities with recorded, read poetry or prose, with just some background noises added, like in rock poetry slams. From another point of reference, it may instead be heard as a somewhat strange rap song with musical accompaniment consisting of a sampled burning match and a second rap voice (“tic-toc”). The context of the performance and listening will determine which generic codes will be applicable, not so much the actual performed sounds themselves.

It is fairly easy to transcribe it in a musical staff, with a pulse, clear rhythms and even melodic lines formed by the interplay of these two voices and the burning match sound. But when it comes to the words, the easiest transcription is to print them alphabetically and add them to the musical transcription in a manner that again installs a bifurcation between the two modes.

You are sitting on a bomb!:

The image shows a musical staff with a treble clef and a tempo marking of ♩ = 80. The staff contains rhythmic notation for the phrase "You are sitting on a bomb!:". The notation consists of eighth and sixteenth notes, some beamed together, and rests. Below the staff, the words "tic" and "toc" are written in a stylized font, corresponding to the rhythmic pulses. The words are arranged as "tic toc tic toc" under the staff. The staff ends with a double bar line.

All these transformations entail decisions as to whether certain sounds are to be shown as belonging to the speech (word) level or to the sound (music) level of the performance. Such decisions can always be challenged: is “tic-toc” a verbal or a non-verbal sound? How is the recital to be denoted and positioned in relation to the staff system?

More questions arise as the second tune, “Rodney King”, makes us understand that this is a record which is consistently built up by sampled voices, noises and instrumental sounds, all electronically processed to create a clearly musical web. The text of “Tictoc” contains intertextual quotations from other texts, songs and films, but they are concealed by the voice of the L.P.C. singer.⁵ In “Rodney King” and most other entries on the album, the sampling techniques make this intertextuality much more obvious. The voices heard are not from the singer who made the first monologue, but “authentic” voices recorded from television programs where they were never intended as musical statements. But here, they seem to get musical qualities. Some are at least as much music as is the vocal line in rap: the sounds and rhythmic organization of the phonemes are carefully utilized and reinterpreted as musical parameters. And in some cases, the human speech is digitally manipulated and given a clearer pitch, so that it sounds as if Rodney King really sings a duet with ex-president George Bush.

One question is who sings here. The *real* Rodney King did not sing but speak those words: it is the King *of this tune* who sings with the President. Still, they are both constructions made and mixed by the L.P.C. collective. None of those real individuals were ever part of the group. A photo in the accompanying CD booklet shows the L.P.C.’s “fabricated” King and President, and many other faces behind the voices sampled on the record, joined in a montage group photo together with the “real” L.P.C. members themselves. This is a pictorial variant of the same collage/sampling technique.

Such authorial queries aside (cf. Fornäs, 1974), the border between song and speech is blurred, and it is hard to decide what is music, what is noise and what is lyrics in this performance. Like with “Tictoc”, if “Rodney King” is at all a song, it forces listeners to reconsider certain conventional distinctions that are routinely made in more traditional pop genres.

L.A. cop: Get up! Get up, sit on the curb!
 Sit on the curb! Hands!
 Stay in the car, stay in the car! [...]

Supporters: Rodney King! Rodney King!

Rodney King: Settin' these fires
President Bush: I felt anger
Rodney King: It's not right
President Bush: I felt pain
Rodney King: It's not gonna change anything
President Bush: I felt anger
Rodney King: It's not right
President Bush: I felt pain [...]



The written text version is here part of what is presented on the disk, which is in its turn a rather condensed version of what is heard. The musical transcription is only a small example of how it could be shown on paper.

It is not hard to find similar examples in popular music, certainly today but also in older genres. Listening for example to Sheila Chandra “speaking in tongues” (on *The Zen Kiss*, 1994) makes it hard to know for sure what is speech and what is song, or exactly what is text and what is music in that song. The use of sampling and other digital technology has not only questioned those concepts of musical work and authorship that were in fact already unstable. By sampling natural sounds, radio talk or television noise, that which used to be heard as non-musical sounds (noise or speech) is creatively woven into a musical (con)text, and its musical qualities like rhythm, timbre or melodic lines are suddenly perceived. Rap has also pointed to ambiguities in the distinction between speech and song, and between text and music, which have actually always existed, but were often naturalized and concealed in traditional rock and pop practices. A rap performance is a speech-song that further destabilizes these boundaries and thus emphasizes that which unites speech and song, thereby more intimately reconnecting the verbal and musical modes.⁶

Words for music

Moving from musical practices to terminology, it is strange to see how vague key concepts like music and text or song and speech actually are. This obscurity might be the reverse of their fruitfully polysemic character: only dead terms can be univocally defined! But their seemingly natural self-evidence conceals the

fact that their boundaries are fleeting and result from multidimensional interconnections of contextualized cultural practices.

One might turn to a dictionary to find valid definitions of *music*, but it is very hard even in estimated musical encyclopedias to find a satisfying discussion of what can be meant by this crucial concept. Most efforts do not to simultaneously manage to satisfy three basic demands: (1) to cover the main genres and forms that are generally experienced and described as music; (2) to distinguish music from speech; (3) to distinguish music from other noises. Many definitions are either too narrow to capture the great wealth of music, independent of time and genre, or too wide to exclude both talk and those natural or technological noises that are not generally thought of as music. I will not here systematically overview all main definitions, and do not long for an all-encompassing definition to solve all problems, but it is interesting to reflect upon what this notorious obscurity implies.

Older, romantic sources think of music as sounds combined in forms of beauty that express emotion. It is in today's late modern period obvious that such definitions fall short of their task. Much music – from Gregorian chant to Glassian minimalism – has no aim and no key effect to express personal emotion. Neither is formal beauty any universally shared musical value, as much film music or thrash metal is deliberately ugly. In these ways such definitions remain too narrow, but they are also too wide, since they cannot satisfactorily distinguish music from speech: public speeches, radio theatre or recited poetry also combines sounds in more or less beautiful forms and with at least some kinds of expressive functions.

Modern versions are less obviously problematic, but similar problems still tend to return in new disguises. The main standard formula today of seeing music as “humanly organized sounds” (e.g. Blacking 3 and 89) needs qualification in order to fulfil its distinguishing tasks. First, it has to be widened or modified, so that the organization of this “sonic order” is understood as being made in the listening process rather than in the physical sound production itself. Blacking (11) is clear on this point, by arguing that music presupposes “the *perception* of order in the realm of sound” (my emphasis). Since one can listen to noises or natural sounds as if they were music, thus making them into music, this human organization of sounds obviously primarily takes place in reception rather than in tone production.

The other problem remains to be solved, however. Which kind of sonic order is being produced? Spoken words and acoustic signals are also humanly arranged sonic organizations, both as they are produced and perceived. To then demand that the sounds must be of a particular kind, for instance composed by notes in a scale or quantifiable rhythms, seems no good solution. Much folk, avantgarde and ambient music does not build upon well-defined single notes or separable rhythms, choosing instead to integrate and recombine a diverse range of more or less strange noises. On the other hand, not only rap, but also everyday speech can with some effort be interpreted as arranged in notes and rhythms. It might be hard to transcribe exact pitches, but again, not very much harder than with micro-intervals and sliding tones in blues or folk ballads, and many spoken syllables can be noted with pretty accurate rhythmic values.

If the sounds in question are instead specified as nonverbal aural expressions, this would seem to exclude songs from being music, since sung words are both verbal and musical at the same time. At least there are then many border cases where sung sound elements can be interpreted both as words and as notes or rhythms. Excluding verbal sounds thus does not help very much, since humanly created non-musical sounds remain, since songs seem to be simultaneously verbal and musical, and since the problem is then only displaced onto the definition of words.

Vocal music is “the encounter between a language and a voice” (Barthes 181). But so is speech, only with a different balance between the two. In speech, language is often so much in focus that the voice/sound aspect – however crucial it may always be for delivering a certain message – becomes secondary. In many singing genres, the voice instead tends to become so important as to push the language (the words of the lyrics and their semantic meaning) into the background.

To define *words* and verbality is hardly easier. If song texts are the words of a song, some of them might be spoken rather than sung, so that the sung words are only a part of the song’s words. (These song words in turn comprise only one level of the whole song text, in the wide sense of this term, including musical, pictorial and gestural elements as well.) Recitative, speech-song and rap blurs this line of division, as does of course the general difficulty to distinguish song from speech. Who and what decides – according to which criteria – which sound elements in a song belong to the verbal level, which sound formations that should be understood as true words? This is always the

result of a contestable and contextualized act of interpretation, interacting with the generic rules upheld by specific interpretive communities. It may often be debated which sounds of a song that can or should be transcribed in/as written words. Whistling, humming or nonsense (e.g. “scat”) singing may either be understood as textless performances where the human voice is just used as a musical instrument, or as creative poetic extensions of the range of verbal expressions, where “A-o-a-a-i” or “M-m-m-m-m-m” would transcribe these emergent, quasi-verbal units. No dictionary can offer any quite clear solution, since words are continuously invented, and nonsensical utterances may actually be heard as meaningful in a given intratextual context.

Is “Be bop-a-lula” a line of lyrics or just an essentially nonverbal transcription of musical sounds that just happen to be produced vocally? “Be bop” can be found in a dictionary but is here used with virtually no reference to its common sense, and even if “a lula” sounds like an English word, it still has no clear meaning. A moaning or a cry may or may not be integrated into a word-based verbal structure. Humming, sighing, shouting and screaming can be heard as parts of pure music, as elements of speech, or just as non-verbal and non-musical noises. Anything that is vocally uttered may in principle enter the realm of verbality and become at least potential words, just like everything in what is socially coded as a musical performance may be heard as music.

Having meaning is not enough to define words, since images and musical themes are also in some sense meaningful.⁷ “Meaning is not inherent in music, but neither is it in language” (McClary 21). “Music is no different from language in that it is a signifying practice with its own particular characteristics” (Shepherd & Wicke 3, also 203). Also, “although music can be textual, it is textual in some very distinctive ways”, and it is important to remember that “popular music’s textuality is comprised of sounds, words, images, and movement” in a rather complex manner (Shepherd 174).

Defining what are the words, lyrics or text in a song may thus be as difficult as to know what precisely is its music. To specify which such sounds are musical (music), which are verbal (speech), which are simultaneously both (song) and which are none (noise) makes necessary a consideration of sociocultural and historically developed contexts, practices and institutions in which these sounds are heard and interpreted. In this sense, David Brackett (125) stresses how power through institutions and discourses ultimately “determines which types of organized sound may be defined as ‘music’.” For

Jacques Attali (4), music is “the organization of noise”, but this particular sonic organization has a specific history which has ascribed to music a particular range of forms, meanings and functions in society. The practices that distinguish music, song, speech and noise from each other, and from non-aural sense modes, are never once and for all given or natural. These concepts are no fixed essences, but the result of complex social and historical discourses and practices, where the interplay between the differential power of individuals, groups and institutions is a crucial determinant.

There is no single, simple way to define the word/music bifurcation. The two terms cannot be defined once and for all, since they are no unitary essences embodied in musical and textual works. There is not even any consistent set of criteria for how to distinguish them. What will be found are contradictory networks of criteria, inherent in shifting practices situated in contexts of a temporal and spatial, historical and social kind. Words and music are open, polysemic and context-bound constructs that always are in process, and can only temporarily and for certain purposes be frozen into well-structured models.

The prevalent terminology for symbolic modes like speech, writing, images and music is often inconsistent, but some preliminary definitions may perhaps be proposed. *Words* are specific combinations of an in principle unlimited number of morphemes (basic meaningful units), in their turn composed of clearly limited sets of letters (in writing) or phonemes (in speech) acknowledged within a certain language. *Verbality* is the symbolic mode that composes utterances, works and discourses by combining such units according to various accepted sets of grammatical, pragmatic and generic rules for their combinations. “*Lyrics*”, like “poetry”, denotes a specific verbal genre – song lyrics are the words of songs – but the term has a somewhat high-cultural aura which is problematic in many popular music genres. Song lyrics may look like poems, but they often work more like dramatic plays.⁸

A special problem is posed by the polysemic character of the term “*text*”. As has already been evident above, much discourse about music uses “text” as almost synonymous with “words” or “lyrics”. On record sleeves, “text & music” is as common as “words & music”, and many of the scholarly quotations mentioned above make similar presuppositions. In literary studies, this may be less problematic since the main texts investigated there are verbal and written, but when music or other means of expression are studied, this usage may cause confusion. A strong tradition of hermeneutic and semiotic cultural studies has

chosen to see all kinds of ordered units in discourse as texts in a much wider sense, going back to the Latin roots of the term “text” as a term for “weaving” or “web”, related also to “texture” and “textile”.

Texts are then at least loosely unified networks of signifying units in any combination of symbolic modes, webs where symbols are intervoven to form meaningful ordered structures. They are complex totalities that are more than the sum of their constituents, in that the textual totality enriches and modifies the significance of each of its symbolic constituents by simultaneously focusing and widening its signifying scope. A document is more than the sum of its letters, words and sentences, just like the painting is more than the sum of its colors and lines or the song “Rodney King” is more than the sum of its many voices, notes and noises. Texts in the widest sense are therefore entities on a higher order than the symbols of which they are composed, and they are to some extent fixated and rounded-off as complex wholes.

In media studies, a text may then be a television program, a poster or a sculpture just as well as it may be a spoken dialogue or a pop song. In this usage, the text of a song includes both words and music, which of course collides with the more common usage of the term. “There are two elements to song, sound and text”, argues Feld (219), in a context which does not make it completely clear if this puts sounds up against words or instead intends a distinction between pure sensual materiality and the generalized orderings of textual structures. I have here tried to stick to the wider sense and preferred using “words” or “lyrics” instead of “texts” when intending the verbal layer of songs. A *verbal text* is then a specific set of words composed as a work into a unit presented and appropriated as a composite whole, independent of their material shape as written, spoken or sung. In a very narrow sense, a (verbal) text is commonly understood as primarily a written or even printed and published verbal work.

Musical terminology has no obvious counterpart to “words” as basic elements of verbal expressions. The term “*music*” denotes the activity and its result as a whole, including either any sounds that are aesthetically organized by listeners, or just those produced by humans and instruments, sometimes even excluding song. Other narrow definitions only include those organized sounds that are marked as music by specific institutions (producers, advertisers, critics), or that fall within the genres one like (“X is not music!”). In many Western genres one can talk about “*tones*” or even “*notes*” as well-defined elements of

fixed duration, pitch and amplitude, comparable to letters (in writing) and phonemes (in speech), but it does not fit rock and pop equally well. “*Sounds*” is an often preferred but not quite sufficient alternative, since sound is also one particular musical parameter and, more importantly, since spoken words are also sounds. Meaningful sound units, parallel to morphemes (situated between phonemes and words) in linguistics, have been named “*musemes*”, a technical term with no evident equivalent in common language.⁹

Words and music

Both words and music can thus be very different things, related to each other in highly divergent ways. The borders between them are notoriously leaky. They cannot be reduced either to a polar opposition or a harmonious complementarity. They are included in cultural spheres where different genres are grouped into institutionalized totalities with many overlaps and blurred margins.

When differences between symbolic modes are thematized, a sharp polarization is often assumed between words and music – or between words and images, but that is another story. These differences are actually neither absolute nor irrelevant, but processual: they are as constructed as are the symbolic modes themselves, and they result from ongoing practices that strive to keep them apart. Music and words are not the two autonomous and opposing fields they often appear to be, but relative gravitational centers within the streams of human communication. They are sometimes seen as being each other’s opposites, but that can only be accomplished by a strongly metaphorical use of key terms that reduces the symbolic modes from the complex totalities which they in practice are into neat inversions or negations of each other, thereby misrepresenting their actual interplay.

Their polysemic character does not imply that they are arbitrary. Within specific interpretive communities and (musical or theoretical) practice fields, they remain pretty self-evident. Experimental genres may question the borders, but most popular music still rests on institutionalized practices where they are clearly separated, for all practical purposes. There are human sets of (listening) activities that concentrate upon the material and formal sound aspects of symbolic webs (while not excluding their meanings), and others that focus on linguistic meaning aspects (where sounds and rhythms are then not absent but

secondary). As long as people thus continue to distinguish between literature (focusing words) and music (focusing sounds), then it will also be relevant to separate musical from verbal aspects of songs, in spite of the virtual impossibility to find any universally valid criterion for this.

The dichotomy of words and music is no true dichotomy at all, but a fragile construction resulting from intersubjective discourses, where each term is constituted more like a *genre* than a physical, objective fact. They are both intersubjective conventions – not objective essences clearly cut off from each other, nor subjectively arbitrary or imaginary illusions. The differences between song and speech or between music and words are neither sharp nor objectively given. Still, in spite of all paradoxical examples, they work pretty smoothly in most aesthetic practices. Each functional, social and generic context creates a kind of interpretive community where it is in a given spatio-temporal and socio-cultural context pretty clear what is what. The arduously accomplished separation of words from music is more of a goal than of a starting-point, rather a tendency than a fact, and it is always threatened by counteracting hybridizing practices. Once the differentiation between words and music has evolved in history and been institutionalized on various levels, symbolic practices cannot avoid relating to it, but its seemingly natural self-evidence is a fiction.

Modern popular music offers plenty of useful illustrations to these general issues, as has hopefully been obvious in this presentation. Even more than for instance in opera, where institutionalized divisions between composers and librettists invites to a more strict separation, many pop songs are conceived and perceived as verbal-musical units, and the experimental genres mentioned above offer plenty of good examples of border-crossings. Song is a hybrid form that *combines* and *is* therefore simultaneously words and music. Songs are more often classified as a kind of music than a type of drama, speech or literature. Perceiving song lyrics apart from music and performance seems like an artificial procedure of translation into another mode or sphere, while humming a song without its lyrics may lessen its impact but remains within the song mode. Song words are thus verbal components of basically musical practices, indicating the overlap between the two modes. And even without the existence of song, spoken as well as written words and live or recorded music would have much in common. They are not completely autonomous or different spheres of communication. Not only do they intertextually interrelate (with various references to and borrowings from each other), but they also share some

important characteristics and are more often inseparably blended than kept rigorously apart.

Notes

¹ This article was originally intended for publication in an anthology on Nordic rock research that regrettably was eventually never realized. It was written in pair with Fornäs (1997), that supplements by scrutinizing some more theoretical aspects of these issues, whereas the intention here is to keep closer to everyday music experiences. A short early version is Fornäs (1998). I wish to thank two anonymous referees for useful comments that helped improving this text considerably.

² For the sake of simplicity, I here concentrate upon the symbolic modes of words and music as organized in relation to the sense of hearing, with only some comparisons with visual representations of such aural forms. I am however well aware that images, kinetics, touch and smell interact with them and should actually be included in a more exhaustive treatment of these issues. Cf. Fornäs (1995) on symbolic modes in general, and Frith (1988 and 1996), Ulf Lindberg or Hillevi Ganetz on lyrics/music-relations in pop and rock.

³ Stephen Feld analyses ways in which bird song (and other natural sounds) are given meaning by human signifying practices.

⁴ Kevin Barry traces a movement towards an emphasis of reception and interpretation in 18th century discourses on music, language and signs.

⁵ Tony Mitchell has pointed out that the words “Tic toc, poor old mom, you are sitting on a bomb” appear in the film *Morgan, A Suitable Case for Treatment* (Karel Reisz, UK 1966), which in its turn is based upon a TV play written by David Mercer.

⁶ The research on noise perception and “soundscapes” mixing musical, verbal and other sounds is informative here; cf. R. Murray Schafer and Helmi Järviluoma (ed.).

⁷ For relevant theories of musical signification, see Nattiez or Shepherd and Wicke.

⁸ “In songs, words are the signs of a voice. A song is always a performance and song words are always spoken out, heard in someone’s accent. Songs are more like plays than poems; song words work as speech and speech acts, bearing meaning not just semantically, but also as structures of sound that are direct signs of emotion and marks of character. Singers use non-verbal as well as verbal devices to make their points...” (Frith, 1988: 120).

⁹ The useful term “museme” derives from Charles Seeger, and was elaborated and refined in the musical semiotics of Philip Tagg. Theo van Leeuwen makes a rather technical effort to systematically “explore the common ground between speech, music and other sounds”.

Works cited

- Attali, Jacques. *Noise: The Political Economy of Music*. Trans. Brian Massumi. Manchester: Manchester UP, 1985.
- Barry, Kevin. *Language, Music and the Sign: A Study in Aesthetics, Poetics and Poetic Practice from Collins to Coleridge*, Cambridge: Cambridge UP, 1987.
- Barthes, Roland. “The Grain of the Voice.” *Image, Music, Text*. London: Fontana, 1977. 179-189.
- Blacking, John. *How Musical is Man?* Seattle / London: U of Washington P, 1973.
- Booth, Mark W. *The Experience of Songs*. New Haven & London: Yale UP, 1981.
- Brackett, David. “Music.” *Key Terms in Popular Music and Culture*. Eds Bruce Horner and Thomas Swiss. Malden, MA / Oxford, UK: Blackwell, 1999. 124-140.
- Cook, Nicholas. *Music, Imagination and Culture*. Oxford: Clarendon Press, 1990.
- Feld, Steven. *Sound and Sentiment: Birds, Weeping, Poetics, and Song in Kaluli Expression*. Philadelphia: U of Pennsylvania P, 1990.
- Fornäs, Johan. “Listen to Your Voice! Authenticity and Reflexivity in Rock, Rap and Techno Music.” *New Formations* 24 (1994): 155-173.
- Fornäs, Johan. *Cultural Theory and Late Modernity*. London: Sage, 1995.
- Fornäs, Johan. “Text and Music Revisited.” *Theory, Culture & Society*, 14:3 (1997): 109-123.

- Fornäs, Johan. "Text + Music." *Music on Show: Issues of Performance*. Eds Tarja Hautamäki and Helmi Järviluoma. Tampere: Department of Folk Tradition, 1998. 104-107.
- Frith, Simon. *Music for Pleasure. Essays in the Sociology of Pop*. Cambridge: Polity Press, 1988.
- Frith, Simon. *Performing Rites: On the Value of Popular Music*, Cambridge, MA: Harvard UP, 1996.
- Ganetz, Hillevi. *Hennes röster. Rocktexter av Turid Lundqvist, Eva Dahlgren och Kajsa Grytt*. [Her voices: Rock lyrics by three Swedish female rock artists.] Stockholm / Stehag: Symposion, 1997.
- Gorbman, Claudia. *Unheard Melodies. Narrative Film Music*. Bloomington / Indianapolis: Indiana UP, 1987.
- Järviluoma, Helmi, ed. *Soundscapes. Essays on Vroom and Moo*. Tampere / Seinäjoki: Department of Folk Tradition / Institute of Rhythm Music, 1994.
- Lanza, Joseph. *Elevator Music. A Surreal History of Muzak, Easy-listening, and Other Moodsong*. New York: Picador USA, 1995.
- Lindberg, Ulf. *Rockens text. Ord, musik och mening*. [The text of rock: Words, music and meaning.] Stockholm / Stehag: Symposion, 1995.
- McClary, Susan. *Feminine Endings: Music, Gender, and Sexuality*. Minnesota / Oxford: U of Minnesota P, 1991.
- Nattiez, Jean-Jacques. *Music and Discourse: Toward a Semiology of Music*. Princeton, NJ: Princeton UP, 1990.
- Schafer, R. Murray. *The Tuning of the World*. New York: Knopf, 1977.
- Seeger, Charles. "On the Moods of a Musical Logic." *Studies in Musicology 1935-1975*. Berkeley: U of California P, 1977.
- Shepherd, John. "Text." *Key Terms in Popular Music and Culture*. Eds Bruce Horner and Thomas Swiss. Malden, MA / Oxford, UK: Blackwell, 1999. 156-177.
- Shepherd, John, and Peter Wicke. *Music and Cultural Theory*, Cambridge: Polity Press, 1977.
- Tagg, Philip. *Kojak – 50 Seconds of Television Music: Toward the Analysis of Affect in Popular Music*. Göteborg: Department of Musicology, 1979.
- Van Leeuwen, Theo. *Speech, music, sound*. London: Macmillan, 1999.

Discography

- L.P.C. *Welcome to Lucky People Center*. Stockholm: MNW, 1993. MNWCD 234.