

SoundEffects



An Interdisciplinary Journal of Sound and Sound Experience

*Anders Bonde
and
Allan Grutt Hansen*

Audio logo recognition, reduced articulation and coding orientation:

**Rudiments of quantitative research integrating branding theory,
social semiotics and music psychology**

*Anders Bonde, Associate Professor
and
Allan Grutt Hansen, Assistant Professor ,
Department of Communication and Psychology,
Aalborg University
abonde@hum.aau.dk
agh@hum.aau.dk*

www.soundeffects.dk



Abstract

In this paper we explore an interdisciplinary theoretical framework for the analysis of corporate audio logos and their effectiveness regarding recognisability and identification. This is done by combining three different academic disciplines: 1) social semiotics, 2) branding theory and 3) music psychology. Admittedly, the idea of integrating sonic semiotics with marketing or branding has been proposed elsewhere (cf. Jekosch, 2005; Arning & Gordon, 2006; Winter, 2011), though it appears novel to apply this cross-disciplinary field from a social-semiotic perspective while, at the same time, focusing on musicological descriptors. We consider as a starting point Kress and Van Leeuwen's (1996, 2006) conceptualisation of 'modality', which is central to their 'visual grammar' theory and subsequently extended to auditory expressions such as spoken language, music and sound effects (Van Leeuwen, 1999). While originally developed on the basis of linguistics and systemic-functional grammar (Halliday, 1978, 1985) and further reinforced by theories of 'intersemiotic translation' (cf. Jakobson, 1959; Eco, 2001) and 'coding orientation' (Bernstein, 1971, 1981), Kress and Van Leeuwen's idea of modality is in this paper connected to notions of brand recognisability and brand identification, thus resulting in the concept of 'Reduced Articulation Form' (RAF). The concept has been tested empirically through a survey of 137 upper secondary school students. On the basis of a conditioning experiment, manipulating five existing audio logos in terms of tempo, rhythm, pitch and timbre, the students filled out a structured questionnaire and assessed at which condition they were able to recognise the logos and the corresponding brands. The results indicated that pitch is a much more recognisable trait than rhythm. Also, while timbre turned out to be a decisive element, RAF did actually cause logo and brand recognition in a substantial way. Finally, there seems to be a connection between the level of melodic distinctiveness and logo and brand recognition. The empirical findings are interpreted and discussed in light of the theoretical framework and the concept of coding orientation.

Introduction

During the last four decades there have been various efforts to bring forth systematic methods to analyse audiovisual products or texts and, specifically, the relation or interaction between their constituent elements, whether labelled 'media', 'modes', 'sensory modalities', 'semiotic systems', 'resources' or, more neutral, 'components' (cf. Constantinou, 2005, pp. 606-611; Bundsgaard, 2005, pp. 54-85; Lauer, 2009, pp. 227-229; Boeriis, 2009, pp. 64-67). The use of terminology depends largely on conventions of the scholarly field in which research on the topic has been conducted, and, accordingly, different categorisation frameworks for qualitative content analysis, combining verbal and non-verbal as well as visual and sonically mediated content or 'semiosis' (Peirce, 1907, p. 411), have appeared. While some contributions have

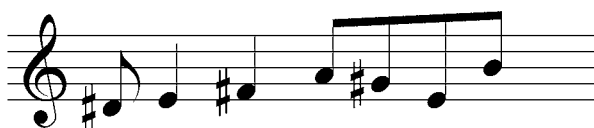
focused on the *separability* or *self-sufficiency* of the semiotic systems, meaning the extent to which, for instance, moving images, verbal messages and music can be grasped and recognised as individual entities (cf. Levinson, 1984; Hoek, 1995; Eck-krammer, 2004; Mitchell, 2005), others have been concerned with *denotational and referential difference*, that is in a Peircian semantic or pragmatic sense (cf. Pauli, 1976, 1981; Barthes, 1977; Tagg, 1979; Chion, 1994; Cook, 1998; Hull & Nelson, 2005). Additionally, and notwithstanding the above-mentioned divergence, there exist a number of academic fields investigating or ‘mapping’ correspondences between the systems: for example, the notions of ‘synaesthesia’ (cf. Cook, 1998, pp. 24f; Van Leeuwen, 1999, p. 144; Haverkamp, 2008) and ‘intersemiotic translation’ or ‘transmutation’ coined by Jakobson (1959, p. 233), each representing ways to conceptualise the transfer of meaning from one system to another while exemplifying a sensuous and linguistic angle, respectively. Intersystemic meaning transfer is, however, a multifaceted matter, but also a complicated and problematic matter. The complications have been demonstrated thoroughly by Eco (2001) in his discussion of ‘translation’ versus ‘interpretation’, such as the transformation of an oil painting into black-and-white print, an object’s interpretation of a nominating verbal expression, or a cinematic adaption of a novel. They are all examples of what he calls ‘intersystemic interpretation’ (Eco, 2001, pp.117-119). As for the problematic aspect, it has often been argued that music does not convey meaning in the same manner as words and images. As phrased by Kivy (1990, p. 174), ‘a piece of music can be mournful but not neurotically mournful over the death of a canary’. In the same way, it is a debated issue whether visual communication can be understood as a kind of ‘language’ on the basis of a ‘grammar’ (cf. Machin, 2007, p. 159f). In both cases, the difficulties are caused by the mutual ‘incommensurabilities’ of the semiotic systems as regards semiotic and expressive potentialities; they may complement each other, but each posits its own identity and, what is most important, they seem reluctant to be compared on the basis of a single unifying measure.¹

The translational or ‘transmutational’ difficulties mentioned above are especially encountered in the field of audio branding and the study of corporate audio logos, aiming continuously to be connected to the visual language of the corresponding brands, their characteristics and symbolic meaning and, accordingly, identified and remembered by customers or, more generally, stakeholders of any kind. Just like visual animated logos, audio logos are acoustic signals of a very short duration (typically between two and six seconds), which may be instrumental or sung melodies of a few notes, but they can also be naturally sampled sounds (vocal or external) and synthetic sound effects. Examples of instrumental melodies are the jingles of the German and Swedish telephone companies T-Mobile (Example 1) and Telia (Example 2), while the classic lion roar of Metro Goldwyn Mayer and

the much newer swoosh sound of Skype represent sampled and synthetic sounds, respectively.



Example 1



Example 2

However, as long-term brand signifiers, audio logos are probably most efficient when based on musical or, more accurately, *melodic* content rather than sound effects, since the latter leads to severe limitations due to the reduced possibilities of how to use them consistently and continually as communicative tools for reaching stakeholders (Bang, 2011, p. 26). Indeed, as emphasised by various theoreticians and practitioners,² audio logos ought to be *flexible*, meaning that they must be operational regardless of musical context and style as well as the technical affordances offered by the specific audio touch points.³ To be sure, the simpler the melody of the logo, the wider may be the ‘palette’ of creative options for its combination or assimilation in textual productions, changing over time and adjusted to different media and campaigns. Additionally, creating an audio logo one must take into account the considerable qualitative differences between, say, an audio logo mediated via multichannel cinema sound and a telephone hold line. Thus, when being flexible, audio logos potentially fulfil desired branding objectives such as *consistency* and *continuity*. Secondly, audio logos must be clearly *distinctive*, in that any audio logo has to be easily distinguishable from other logos. This being so, they should, thirdly, be *identifiable*, *recognisable*, *memorable* and linked to a specific brand while, fourthly, they have to be *suitable* (‘fitting’ or ‘congruent’) with regard to the attributes and values of the corporate brand.

Now, considering the above-mentioned branding guidelines, one might pose the question whether there are any musical solutions that meet them all at once. To begin somewhere, initially, we explore an interdisciplinary theoretical framework for the analysis of corporate audio logos and the investigation of their effectiveness regarding logo and brand recognisability as well as brand identification. This is done by combining notions of brand recognisability and brand identification with

social semiotics and music psychology. Whereas the idea of integrating sonic semiotics with marketing and branding has been proposed elsewhere (cf. Jekosch, 2005; Arning & Gordon, 2006; Winter, 2011), it appears novel to apply this cross-disciplinary field from a social semiotic perspective while, at the same time, focusing on musicological descriptors as a basis for empirical audience research. Starting with a consideration of the terms ‘modality cue’ and ‘coding orientation’ as classified by Kress and Van Leeuwen (1996) and Van Leeuwen (1999), we will discuss a theoretical concept aiming to identify the ‘essence’ of any audio logo containing melodic material, and we will examine its empirical testing through an experimental-psychological conditioning experiment in which upper secondary school students assessed the recognisability of five existing audio logos which had been manipulated in regard to fixed design characteristics such as tempo, rhythm, pitch and timbre.⁴ The empirical results are finally evaluated in light of the theoretical concept and social semiotics and in particular coding orientation.

Modality cues and coding orientation

Despite the intersystemic complexities and problems as proposed by Jakobson, Eco and others, a subbranch of semiotics labelled ‘social semiotics’ was developed during the late 1980s on the basis of the systemic-functional grammar theory of Halliday (1978, 1985) with the purpose of establishing an analogous visual approach while maintaining strong emphasis on the significance of sociocultural context as a critical determinant for making meaning (cf. Hodge & Kress, 1988; Thibault, 1991; O’Toole, 1994; Kress & Van Leeuwen, 1996). Taking the three communicative (ideational, interpersonal and textual) metafunctions as a point of departure for answering questions such as ‘How do images relate the people, places and things they portray to each other, so as to form coherent representations?’ and ‘What kind of “interpersonal resources” do the images use to create a relation between the image and the viewer?’ (Van Leeuwen, 1999, p. 190), Kress and Van Leeuwen (1996) developed an analytic-descriptive system for characterising the meaning potential of any given image (i.e. its context-dependent semiotic properties). They defined eight visual modality ‘markers’ or ‘cues’ (1996, pp. 165-167), each expressed as continuous variables or parameters with an infinite number of articulatory levels between two opposite extremes, a minimum and a maximum (cf. Fig. 1). The principal idea behind the grammar is ‘that modality is realized by a complex interplay of visual cues’ (Kress & Van Leeuwen, 1996, p. 167), each independently referring to the amount of *reliability* assigned to the image; and by striving for a consistent approach, the authors seem to comply with Hjelmslev’s ‘empirical principle’, stating that ‘the description shall be free of contradiction (self-consistent), exhaustive, and as simple as possible’ (1961, p. 11). Indeed, by consistently using a unifying epistemological

concept, Kress and Van Leeuwen offered an all-inclusive, but transparent and logical framework for mapping out the socially constructed resources for the communication of meaning through images. Even though they do not have the same semiotic accuracy as language, the modality cues make up a powerful toolbox ‘for allowing us to think about the details of visual communication, for systematically analysing the use of visual semiotic resources’, as noted by Machin (2007, p. 160).

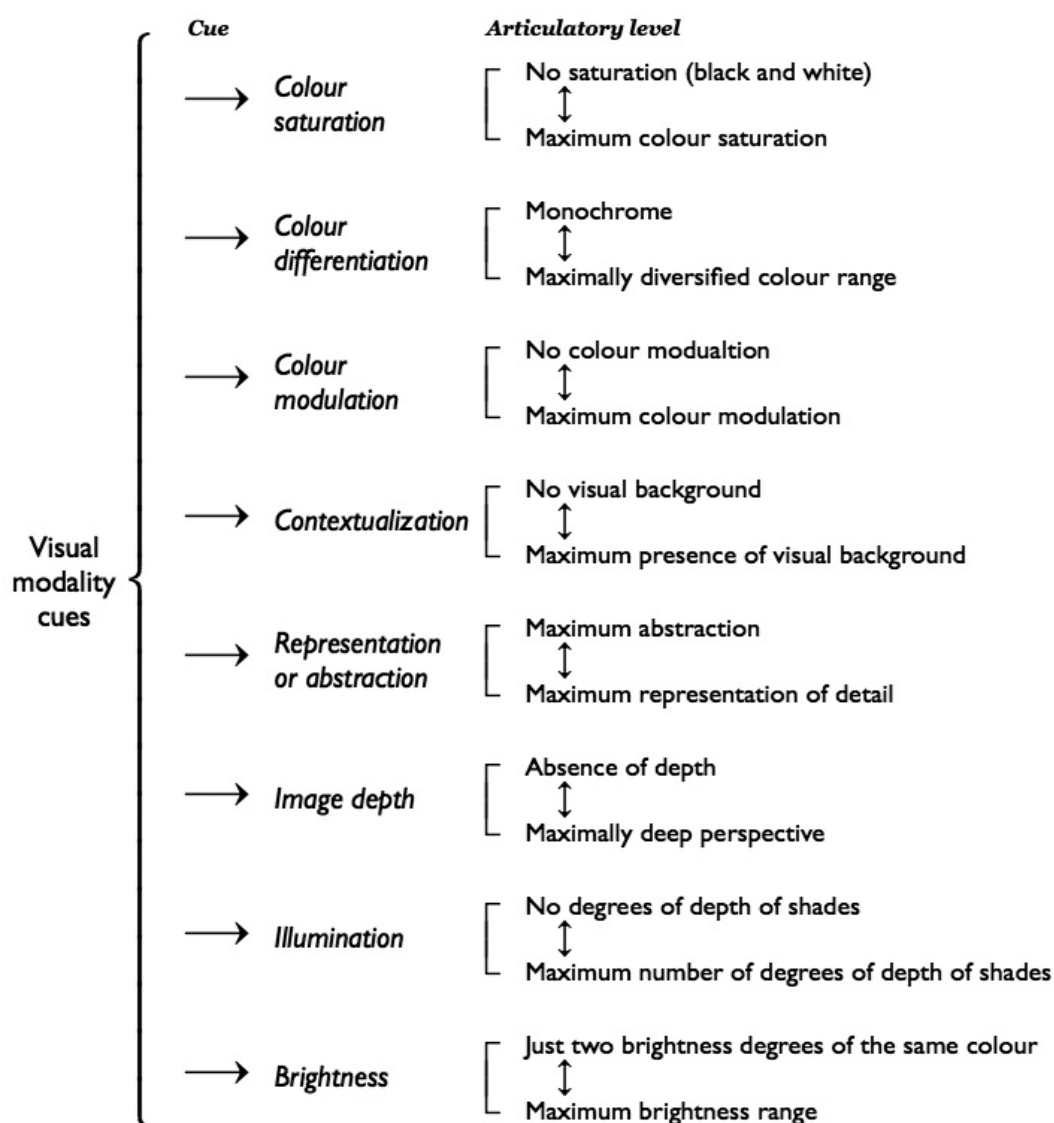


Fig. 1

A few years after Kress and Van Leeuwen's joint work, Van Leeuwen went further and offered a 'sonic grammar', asserting that '[t]he modality of sound can be approached along the same lines as the modality of images' (1999, p. 170), and so he put forward a

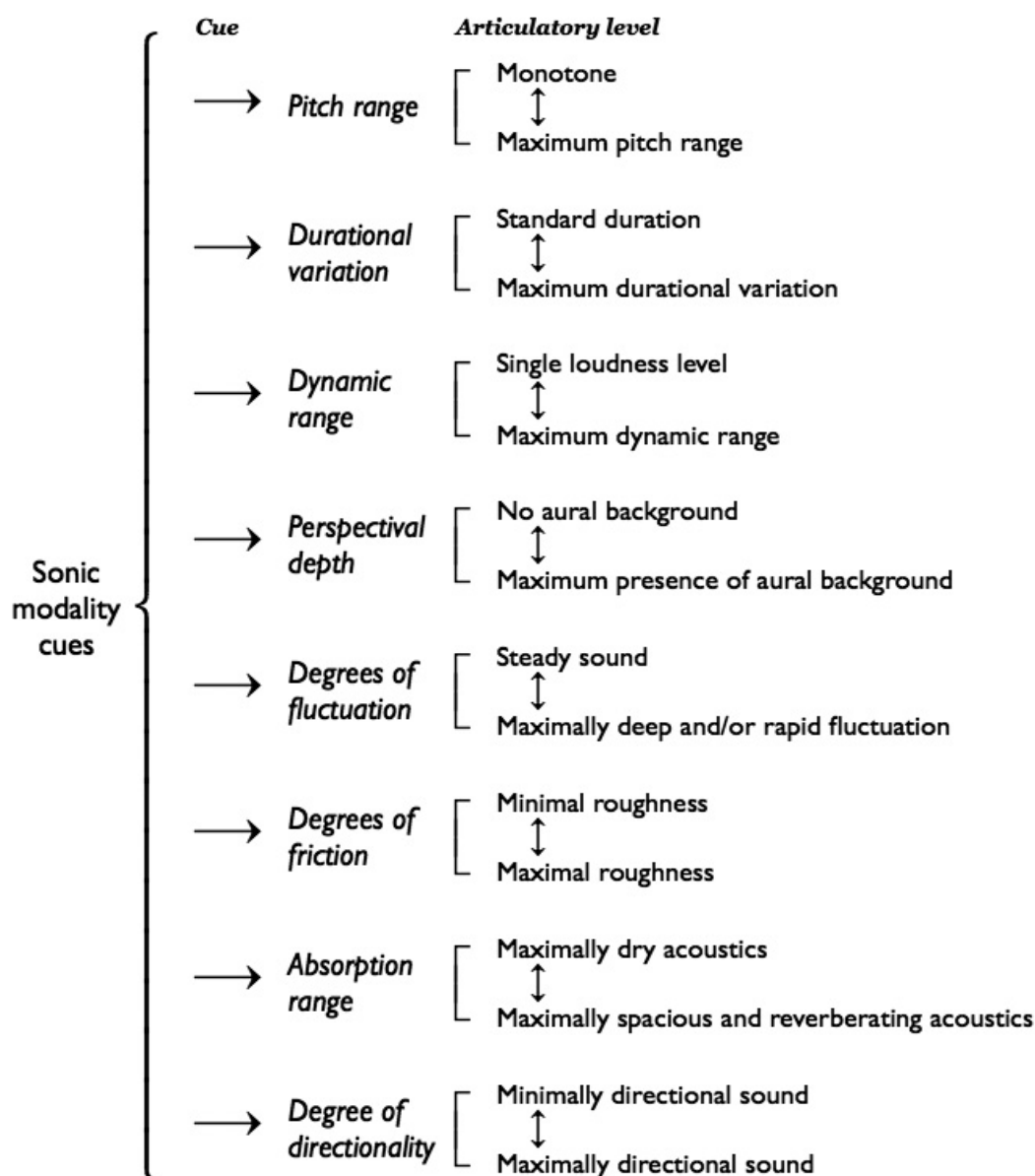


Fig. 2

list of 'sound modality cues', defined as manipulable 'articulatory parameters' (1999, p. 172) for auditory expressions such as spoken language, music and sound effects, also reaching the number of eight and similarly expressed as continua between sets of binary oppositions (cf. Fig. 2).⁵ Due to fundamental phenomenological differences between images and sounds, concerning how the metafunctional resources are configured (though see Wingstedt et al., 2010, p. 196), Van Leeuwen's approach was initiated by a focus on *presentation*, that is, what people can do with sounds

(e.g. voice quality or phonetics in speech and timbre in music), rather than *representation* (i.e. how sounds communicate or ‘mean’ potentially and symbolically). To use Van Leeuwen’s own phrase, ‘In the semiotics of sound you cannot represent “disharmony” without actually having two “voices” (human, instrumental or otherwise) clash with each other. Nor can you represent ‘tenderness’ without actually addressing your listener in a tender way’ (1999, p. 191). Nonetheless, in spite of such differences, all 16 modality cues, being visual or sonic, relate to ‘material aspects’ or the ‘substance’ of a given analytical object, and in that sense the visual and sonic grammars turn into mutually correlatable systems, each including a definite set of semiotic resources for the construction of meaning, all dependent on social and cultural context.⁶

The latter, regarding the notion of *context*, holds crucial importance, not at least in a social semiotic sense, seeing that the evaluation of ‘modality’ – being a measure of reliability, as noted above, that is, how true, accurate or plausible a representation (verbal, visual or audible) is to be considered – depends on ‘habitus’ (Bourdieu, 1977, 1986) or, rather, ‘coding orientation’, a term originally coined by Bernstein (1971, p. 135). It means that thinking, watching, listening, communicating and interacting are guided by discursive and socially established practices, which ‘regulates orientation to meanings [...] through *selection* [of] specific textual productions’ (Bernstein, 1981, p. 329; emphasis in original). In other words, there are as many different realities as there are social groups. Indeed, as stated by Kress and Van Leeuwen,

[A] social semiotic theory of truth cannot claim to establish the absolute truth or untruth of representations. It can only show whether a given “proposition” (visual, verbal or otherwise) is represented as true or not. From the point of view of social semiotics, truth is a construct of semiosis, and as such the truth of a particular social group, arising from the values and beliefs of that group. (1996, p. 159)

However, coding orientation is further constrained by the *situational context* in question (Adlam, 1977, pp. 30-31),⁷ and this view is also supported by Kress and Van Leeuwen (1996). In this way, they propose a typology distinguishing four visual coding orientations (1996, pp. 170-171), which are partially resumed by Van Leeuwen (1999, pp. 159-163) and furthermore adapted to sound (1999, pp. 177-180). In the following paragraph we shall recapitulate the four types, one after another, given that they help shed light on the extent to which the theoretical concept behind the recognition experiment has empirical validity.

The *first* type of coding orientation, a ‘technological’ one, refers to ‘the practical usefulness of the image’ (Van Leeuwen, 1999, p. 161). Here, modality tends to *increase* the more an image appears instructive for taking action of some kind, for instance, how effectively an architectural drawing reveals the thinking of its creator. Accordingly, levels of image depth (perspective), contextualisation (background) or illu-

mination (shading) are normally reduced to a minimum. Generally, the fewer the articulatory details, the higher the modality of the image, and therefore colours are usually absent, unless, of course, they serve as visual aid to differentiate between important details.⁸ A *second* type, marked as 'sensory', is characterised by almost the opposite features, since maximally saturated, modulated and differentiated colours are constituents of high modality. The reasoning is that in printed food advertisements or in pictorial art, such as the flower paintings of Georgia O'Keeffe (1887-1986), affective response (arousal) and pleasure are highly prioritised. It is the same with sound, where an amplification of perspectival depth, absorption, fluctuation and friction most likely leads to increased emotional effects. As a *third* possibility, the authors mention 'abstract' coding orientation, which is similar to technological coding orientation as regards the prioritisation of reduced articulation. Here, 'modality is higher the more an image reduces the individual to the general, and the concrete to its essential qualities' (Kress & Van Leeuwen, 1996, p. 170). There are, however, prominent differences too, for example in terms of colour saturation in which the modality of supernatural colouring exceeds the modality of ordinary colour photographs. This holds true especially in modern art, looking beyond the surface for 'deeper realities' or seeking 'to redefine reality' (Kress & Van Leeuwen, 1996, p. 171), and in that sense colours are valuable. According to Van Leeuwen, it is problematic to think of abstract coding orientation in the realm of music, because 'music is always also sensory' (1999, p. 178). Even when being represented as conventional staff notation, leaving out information of timbral features and sound location (including e.g. friction, fluctuation, perspectival depth, reverberation and directionality), pitch range and durational variation 'may still be exceptionally wide, and hence exceptionally charged with emotion' (Van Leeuwen, 1999, p. 178). This is a significant point that we shall follow up in the next section in which we elaborate further on the analogy between visual and sonic cues. The *fourth* and last type of coding orientation recognised by Kress and Van Leeuwen is labelled 'naturalistic', being a measure of how close a visual or audible representation is to what counts as 'real', 'normal' or 'everyday' (Van Leeuwen, 1999, p. 160, 179). While the technological, sensory and abstract coding orientations can be comprehended as markers of 'particular institutionalized communities' (Zappavigna, 2005, p. 513), naturalistic coding orientation seems to be the 'default' and dominant type, regardless of social class, education or habitus.

To sum up, the theoretical contributions of Kress and Van Leeuwen (1996) and Van Leeuwen (1999) pave the way for examining intersemiotic interaction in general, seeing that the modality cues are representative of generic attributes with potential for affective and emotional response as well as symbolic meaning-making. Moreover, what counts as high or low modality depends on coding orientation, which is a matter of discourse, situation and sociocultural habitus. We believe that

such insights can be profitable applied to studies of music and sound in marketing communication, advertising and branding,⁹ which have traditionally been confined to, on the one hand, media researchers and musicologists, approaching qualitative and hermeneutic methods for a historical survey or the interpretation of aesthetic and social values, and, on the other hand, advertising and market researchers, preferring quantitative and experimental methods (Graakjær & Jantzen, 2009, p. 15). Thus, with the aim of filling this ‘gap’, we will now describe how selected sound modality cues make up the necessary conditions for creating flexible and distinctive audio logos.

A reduced articulation form (RAF)

As implied above, the valuable attribute of the social semiotic approach of Van Leeuwen (1999) is that it includes, to begin with, a straightforward, well-defined, transparent and generalisable model for the systematisation of the articulatory complexity of sounds (cf. Fig. 2). Indeed, pointing out a limited number of descriptors (cues) makes the complexity more manageable, even though the possibilities of combining different levels (from minimum to maximum) of pitch range, durational variation, dynamic range, perspectival depth etc. are infinite. Additionally, integrating this model with the typology of coding orientations provides a means of evaluating, on a reasonably consistent basis, a sound’s level of modality, dependent on the specific configuration. What is interesting is that the approach represents certain possibilities in the field of audio branding, which, to our knowledge, have not been systematically addressed until the study of Bang (2011), the principal conceptual idea of which we shall account for here.

Following Van Leeuwen (1999), it seems safe to say that the more a sound’s articulatory complexity is reduced, the higher the level of abstraction at which the sound is encoded, that is, the less naturalistic or ‘real’ the sound will appear. This is similar to studies of visual shape and texture, where the modality configurations of, say, a high-resolution depth-of-field photography and a newspaper cartoon represent two opposites, encouraging natural and abstract coding orientation, respectively. However, even the most primitive or simplistic kind of drawing or, for instance, a pictogram may represent a rich semantic expression that very much conveys the essential meaning of the depicted subject, object or phenomenon. The same goes with animated motion films featuring geometric shapes, such as the one used in an experimental study of Heider and Simmel (1944), in which a large triangle, a small triangle and a circle move around an empty square, creating an illusion of interpersonal interaction (e.g. chasing, fighting and hiding), like three human characters with intentions, emotions and personalities. Correspondingly, we propose that the sonic complexity of musical signals such as audio logos can be reduced and still

meet the critical needs of brand recognition and brand identification. In the following we shall designate this format as 'RAF', meaning Reduced Articulation Form. The point of making this audiovisual analogy is not to claim that the perceptions of images and sounds (or music) are 'alike' and comparable in every respect (cf. the discussion above on semiotic incommensurability), but concerning abstraction level there are some prominent parallels between shaping a pictographic image of dots and lines and shaping a melody of sine tones. Both expression forms are figurative, though highly abstract due to a minimised level of articulation complexity. Still, having said so, the analogy has its limits or flaws. Compared to a pictogram with minimal articulation in all visual modality cues, which is nevertheless figurative, RAF implies different pitches and durational variation (rhythmic characteristics) to be identifiable and recognisable. However, RAF is nonetheless a highly reduced expression form, and if such combination of high abstraction and 'semantic richness' can be applied on audio logos, the aforementioned branding guidelines of distinctiveness and flexibility are met. Thus, leaving all the articulatory parameters on a minimum, except for pitch range and durational variation, it should still be possible to create melodically *distinctive* audio logos, which at the same time are *flexible* enough to be transmitted through virtually all kinds of audio touch points and to be merged in any musical style or genre, adapted to the specific culture, market segment and the campaign in question. The rationale is then that by fulfilling the criterion of flexibility the matters of *consistency* and *continuity* are practicable too. Actually, the audio logo of McDonald's has represented a profound example of flexibility, consistency and continuity ever since the company licensed the Justin Timberlake single *I'm lovin' it* in 2003.¹⁰

Now, while being flexible for sure, one may ask whether the RAF concept has the potential to fulfil the requirement of logo and brand recognisability together with brand identification. There are certainly reasons to believe that it *has*, seeing that a sequence of musical tones, while being perceived together as a unified entity (according to common criteria of Gestalt perceptual grouping), represents a dynamic and figurative kind of expression in any case. Given a sufficient variability of pitch intervals and durations, making the tone sequence distinctive from other tone sequences or melodies, it should also be recognisable once learned. One could perhaps even probe whether it might be easier to recognise melodies when timbre is left out, since they appear more clear-cut in their simplest form. Such argument does not come out of thin air; in fact, there is empirical evidence of such matter in the visual domain. According to Chandler (2013), when comparing 'a photographic image with a cartoon image of the same thing, the photograph is likely to be judged as more "realistic"; the mental schemata involved in visual recognition may be closer to the stereotypical simplicity of cartoon images than to photographs'. Chandler refers to an experiment by Ryan and Schwartz (1956), showing that a car-

toon drawing of a finger snapping hand was identified quicker than a photograph of the same image, which, he concludes, ‘underlines the importance of perceptual codes in constructing reality’ (Chandler, 2013); that is, the results point towards the fact that the drawing under the given circumstances, according to the specific coding orientation of the subjects in the study, was perceived as more real or, rather, more accurate or plausible than the photograph. To support his argument further, Chandler cites Eco (1976), who, from a theoretical perspective, stresses the linking between iconic depiction and cultural convention. We shall here cite Eco at length:

Maybe an “iconic” solution is not conventional when it is proposed, but it becomes so step by step, the more its addressee becomes acquainted with it. At a certain point the iconic representation, however stylized it may be, appears to be more true than the real experience, and people begin to look at things through the glasses of iconic convention. (Eco, 1976, pp. 204-205)

However, the fact remains that the iconic depiction must be clearly *distinct* to be easily recognised as described above, and, as a matter of course, that also goes with audio logos. The important question is, then, how melodic distinctiveness, in terms of pitch range and durational variation (Van Leeuwen, 1999), relates to logo and brand recognisability when focusing on the RAF concept.

A mixed-method study: Design and results

To explore these matters, an experimental reception study was conducted among 137 students in seven classes in two upper secondary schools in Northern Denmark (cf. Bang, 2011). The aim of the study was more specifically to investigate in general the recognisability and identification of audio logos as RAF vis-à-vis audio logos in their original form and a possible relation to melodic distinctiveness. An intermediate aim was to explore whether pitch range or durational variation represented the most important modality cue. For this reason, a mixed-method approach was used, which included both quantitative data from recognition and identification tests and quantitative as well as qualitative data from content analysis focusing on melodic distinctiveness. Audio logos of five corporations were chosen, which included Intel (Example 3), McDonald’s (Example 4), Nokia (Example 5), JYSK (Example 6) and Elgiganten (Example 7), all considered well known at the time of research (February, 2011), as each of them had been transmitted on national television for at least eight years.¹¹ Even with reduced articulation in accordance with RAF, they differentiate significantly from each other on several design characteristics (cf. Table 1). Apart from different origins (some of them derived from pre-existing musical material, and some designed specifically for the companies in question), the logos consist of a varying number of pitches (from five to 13) and structural units (from one to

three). They also diverge in terms of key and resolution. For instance, while the JYSK logo, as the only one of the five, is open-ended, concluding on the second step of the C-major scale, the other logos are more melodically closed, ending on the first step, or, as regards the Intel logo, the fifth step (though with a clear tonal finalisation). Besides, the melodic contour of each logo is unique due to varying combinations of pitch and duration content as well as a varied number of half-tone steps between highest and lowest pitch (i.e. pitch range or ambitus). All such differences (and others) influence how distinct the logos might be characterised. However, without getting into detailed musical analyses of each logo, it seems reasonable to assert that the quantitative content data (regarding number of pitches, ambitus, interval direction changes, durational changes and structural units) are crucial elements and, consequently, it is clear that the Nokia, McDonald's and JYSK logos are superior in that respect.











Logo	Origin	Pitches	Pitch contour	Ambi- tus	Interval- direction changes	Resolution	Duration contour	Durational changes	Structural units
Intel	Specially designed.	5		7	3	Closed		1	2(1)
McDonald's	Pre- existing.	9		9	4	Closed		4	2
Nokia	Pre- existing.	13		15	6	Closed		5	3
JYSK	Pre- existing.	7		12	3	Open		4	1
Elgiganten	Specially designed.	9		5	4	Closed		2	2

Table 1

In the experiment the students were divided into two groups comprising 66 and 71 subjects, respectively, with each group participating in two different test conditions labelled 'Test 1' and 'Test 2' (Table 2). In both test conditions all five audio logos were played in four different versions, one after another, where tempo and articulatory complexity were independent variables. Put in another way, the experiment was carried out as two separate musical commutation tests focusing on pitch and rhythm, respectively, beginning with the most reduced and therefore most unrecognisable version and ending with the original audio logo.¹² The subjects received a questionnaire and were told to mark with an 'X' one of the four versions after which

<div>Independent variables</div> <div>Tempo</div> <div>Articulatory complexity</div>		<i>Test 1: Pitch (n=66)</i>	<i>Test 2: Rhythm (n=71)</i>
	<i>I</i>	Pitches only, slow tempo	Rhythm only, half tempo
	<i>II</i>	Pitches only, fast tempo	Rhythm only, original tempo
	<i>III</i>	Pitch and rhythm, original tempo (RAF)	
	<i>IV</i>	Original sonic logo	

Table 2

I recognized the sonic logo	I <input type="checkbox"/>	II <input type="checkbox"/>	III <input type="checkbox"/>	IV <input type="checkbox"/>
I recognized the brand	I <input type="checkbox"/>	II <input type="checkbox"/>	III <input type="checkbox"/>	IV <input type="checkbox"/>
Brand name	<hr/>			

Table 3

they believed to recognise the audio logo – and another ‘X’ when recognising the brand. Finally, they were told to identify the brand by writing the name (Table 3).

Data from the study have been presented and analysed elsewhere (Bang, 2011; Bonde & Bang, 2012; Bonde, 2013). In the present paper, we have prioritised examining the following questions:

1. In terms of logo and brand recognition, how effective is RAF compared to audio logos in their original sounding form?
2. To what extent is logo and brand recognition along with brand identification dependent on pitch content or rhythmical design?
3. To what extent is brand recognition coherent with brand identification?
4. Do the responses suggest any correlations between recognition on the basis of RAF and melodic distinctiveness?

When analysing the results of the five logos in total, comparing average recognition and identification rates on the basis of the four subsequent logo versions, one might come to a number of overall findings and conclusions (cf. Fig. 3). To begin with, there is a significantly larger percentage share of subjects recognising the logo than recognising the brand, and a smaller share of subjects ending up identifying or nearly identifying the correct brand than reporting brand recognition after

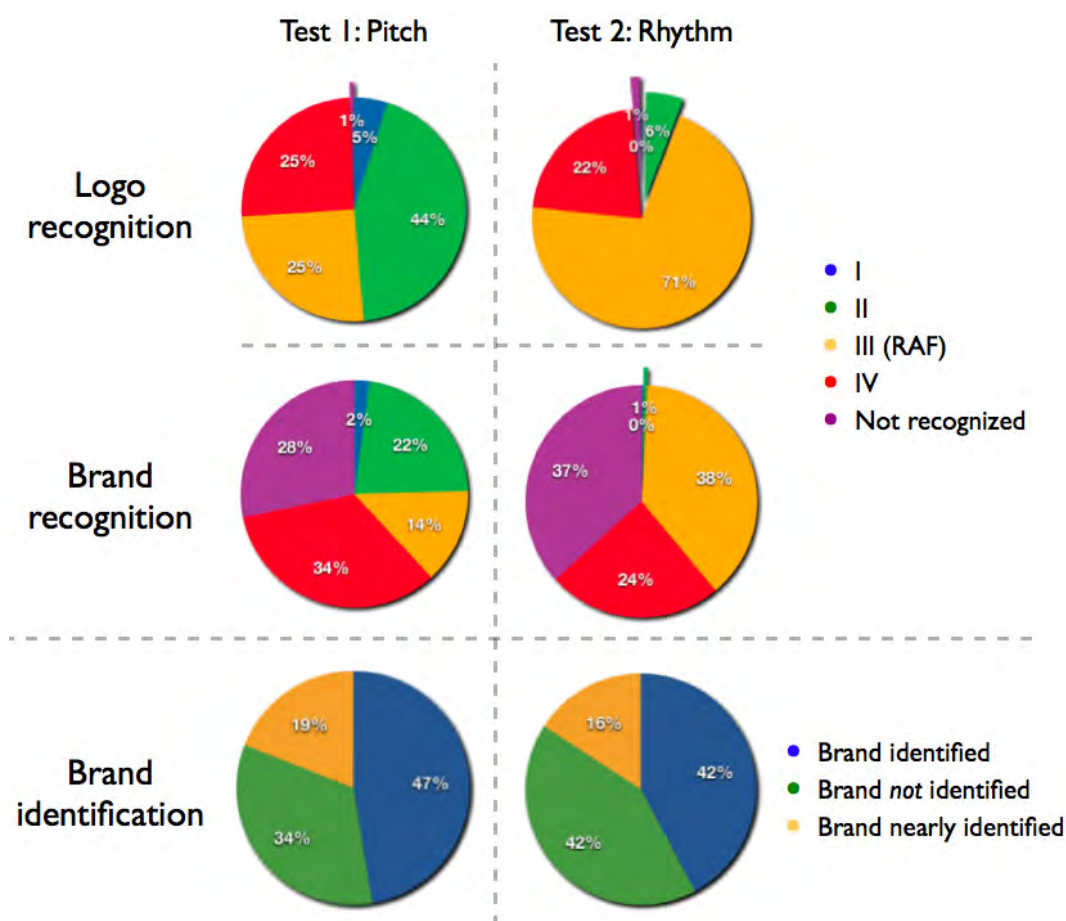


Fig. 3

version IV. This is hardly surprising, considering a signifier might very well evoke recognition in people who are unable to link to a particular signified; though, as we shall see, there are considerable differences between the five logos. As for now, we will just conclude that brand recognition and brand identification are different matters. We may also note that the logo recognition after playing version IV is almost 100 per cent, which confirms fully our notion that all five logos have been largely exposed during the past eight years, and it explains why the corresponding brand recognition is also relatively high.

Secondly, it is clear that pitch is a far more significant parameter than rhythm for recognising audio logos and corresponding brands. In that respect, compare the red colours in the left and right columns (conforming to Test 1 and 2, respectively), and notice the overwhelming percentage difference in recognition increase after playing version II. Furthermore, it appears that a normalised tempo in Test 1 (regarding pitch) has a tremendous effect on logo and brand recognition, even without rhythmic information, whereas, on the contrary, changing the rhythm from half time to

original time (in Test 2) has almost no effect at all.¹³ Also, in terms of brand identification, it is clear that pitch is more effective than rhythm, since more subjects have identified the exact name of the brand or a comparable brand name. The study does not reveal why this is so, but one possible reason could be that rhythmic perception generally requires auditory stimuli with durations longer than a few seconds. One could also question whether the results would be the same if the tests were conducted in another discourse or sociocultural field with other participating subjects with, perhaps, different coding orientations, but that remains pure speculation. However that may be, the findings might be an important lesson for marketers and advertisers to keep in mind when pursuing an identifiable and recognisable brand over time, though we shall return to this discussion in the final section of the paper.

As a third point, it is clear that recognition increase, after playing version III (RAF), equalises the percentage differences in the two test conditions. Thus, compare the combination of the blue, red and green colours in the left and right columns,¹⁴ and notice the percentage similarities when adding together the shares. Actually, in Test 2 (regarding rhythm) there seems to be a larger effect of RAF resulting in logo and brand recognition rates which altogether exceed the corresponding recognition rates in Test 1. This may perhaps come as a surprise, though one explanation could have to do with the so-called 'aha experiences' (Eureka effect), emerging suddenly after not recognising the preceding two stimuli. We shall return to that issue in the discussion. Whatever the circumstances, in addressing our first question we conclude that RAF *does* cause recognition in a substantial way, even though timbre appears to be a decisive element, especially in Test 1 where the increase of logo and brand recognition after version IV (purple colour) is considerable. As for the latter, this can, admittedly, be interpreted as if RAF is *not* to be considered in itself sufficiently effective in causing logo and brand recognition (cf. Bonde, 2013). However, the *raison d'être* is as follows: In combination with pitches in normalised tempo, RAF apparently synergises an even higher recognition rate after playing version IV than it does in combination with rhythm in original tempo in Test 2. Put in another way, the higher the percentage recognition rate after version II (red colour) in proportion to the percentage recognition rate after RAF (green colour), the higher the percentage recognition rate after version IV (purple colour). The effect of the conditional difference is present for both logo and brand recognition, but as regards the latter, the effect after version IV is significantly higher (nine per cent), and that indicates a further synergising effect. Indeed, it seems that the synergised logo recognition effect has a positive impact on brand recognition, but only in Test 1, since pitch recognition is a crucial factor. Therefore, taking nothing else into account, the recognition rate after RAF is almost the same in both test conditions, but the synergy effect might be considerably larger in Test 1.

Now, since the diagrams in Figure 3 are based on all five logos in total and, as such, represent an average result, it might be illuminating to check whether there are structural coincidences or dissimilarities when comparing with the results on basis of the individual five logos. In Figure 4 we provide an overview, and, as can be seen, despite striking differences, we may also notice important structures correlating with the overall result. Taking the latter first, it is clear that pitch, once again, causes a much higher recognition effect than rhythm, and the above-mentioned synergising effects in Test 1 seem to be present too. For instance, when comparing brand recognition in the two test conditions, the percentage shares of 'Not recognised' (orange colour), as can be seen in the cases of the Intel, JYSK and Elgiganten logos, are considerably smaller in Test 1. Moreover, in Test 2 RAF leads to an overall recognition rate that either surpasses or is similar to the corresponding recognition rate in Test 1 (except for brand recognition in the case of JYSK). As for the individual differences, it is perhaps surprising that only very few students were able to identify correctly the Intel brand, while about half (averaging the two test condition results) wrote alternative, but nearly equivalent nouns or brand names such as 'computer', 'Dell', 'Microsoft' and 'Windows'. It is, for that matter, interesting and rather paradoxical, then, that the 'Intel Inside' sound is widely renowned as a model example of effective audio branding (cf. Jackson, 2003, pp. 2-3; Bronner, 2004, pp. 51-52; Lindstrom, 2005, p. 22; Roth, 2005, p. 25; Straka, 2007, p. 67; Kusatz, 2007, p. 3). Thus, even though the Intel logo was recognised after version IV by all students (in both tests), it appears that Intel's partnership with Dell and other computer companies has resulted in brand confusion.¹⁵ Also in the case of Elgiganten, more students have made alternative matches (e.g. 'electronics' and names of competing retail chains) than the correct match.

The results regarding the Intel logo differ in more than this respect. One may, for instance, notice the poor logo and brand recognition rates after RAF, unlike the

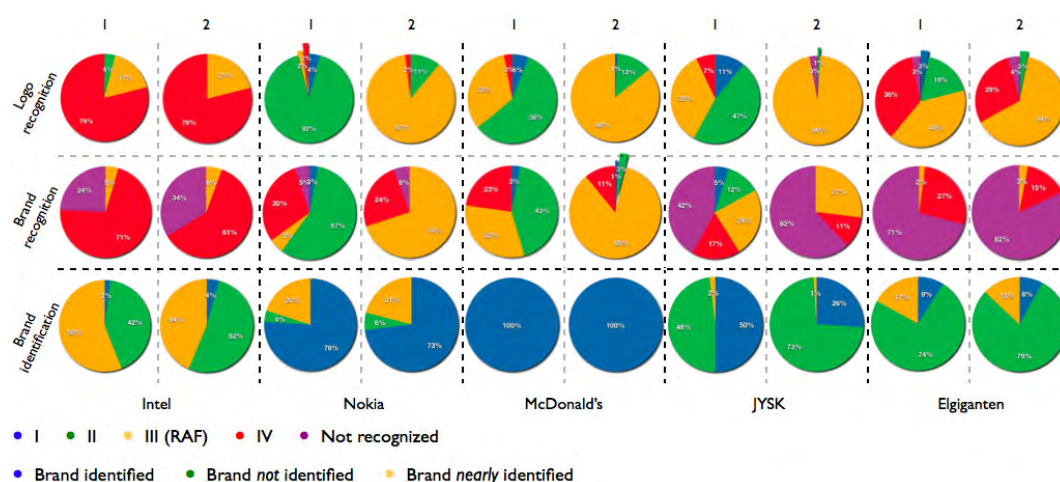


Fig. 4

cases of McDonald's, Nokia and JYSK in which logo and brand (also after RAF) were much more recognised. (Only in the case of Elgiganten, the percentage is lower.) A comparably low level of melodic distinctiveness might explain this, such as few pitches, few interval direction changes, few durational changes, a moderate pitch range (ambitus) and just a single structural unit (cf. Table 1).¹⁶ Nevertheless, it is not clear if these particular features are caused by correlations between, on the one hand, recognition on the basis of RAF and, on the other hand, melodic distinctiveness. Admittedly, the most recognised logos after RAF are the McDonald's, Nokia and JYSK logos, and they are also the ones with the largest pitch ranges (ambitus) and the highest number of durational changes. However, as for the number of pitches, interval direction changes and structural unites, the Elgiganten logo exceeds the JYSK logo. At any rate, it is worth noting that the results on the basis of the individual logos might very well be due to factors not considered in the experiment, such as the subjects' previous experience and acquaintance with the logos and brands. For that matter, it appears that the three well-recognised logos are derived from pre-composed or pre-existing music material,¹⁷ while the two remaining logos are specifically made for the brand, and that, of course, might have an important influence on the results, but we cannot know for sure.

Summation and discussion

It is an underlying premise in this paper that RAF fulfils the criterion of musical and technical flexibility, allows for consistency as well as continuity and provides the opportunity of (melodic) distinctiveness, given the numerous combinatory possibilities in the domain of pitch range and durational variation (or rhythm), corresponding to the first and second modality cues of Van Leeuwen (1999). What remains to be tested, then, is the effectiveness of RAF regarding logo and brand recognisability, brand identification as well as brand (or product) fit, though only the former two issues have been considered in the present experiment. When recapitulating one by one the four exploratory questions posed above, we offer the following answers:

Firstly, *the effect of RAF is substantial*, but also subtle due to the experimental procedure (cf. Table 2) in which self-reported recognition is informed not only by listening to the present logo version, but also through remembrance of the previous version(s). This is why we can speak of a *synergising effect*, meaning that RAF (green colour) stimulates, in combination with version II (red colour) and version I (blue colour), extensive logo and brand recognition after version IV (purple colour). However, this effect is present only in Test 1.

This clearly points towards the fact that *pitch is a much more decisive feature than rhythm*, at least in this experiment. The huge percentage differences between the test conditions after version II (red colour) show evidence of that, and it appears

that hearing pitches only instead of rhythm only (corresponding to versions I and II in Tests 1 and 2, respectively) also has a positive effect on brand recognition and brand identification, seeing that the share of students ending up *not recognising* (orange colour) and *not identifying* (yellow colour) anything is significantly higher in Test 2. Likewise, we find that the suggested possible Eureka effect in Test 2 exemplifies the ‘power’ of pitch; when added to already perceived rhythmic stimuli, recognition ‘explodes’ (i.e. multiplies manifold times), while rhythmic information in Test 1, on the contrary, contributes to a modest increase of recognition (57 and 59 per cent for logo and brand, respectively), resulting in an overall slightly lower share of students reporting recognition after RAF. Viewed from a musicological perspective, such results are quite interesting and suggest similar empirical psychological investigations of, say, transformation and development of musical ideas in classical compositions; but also in the field of marketing, advertising and branding, there are some important perspectives to consider. For instance, assuming that marketers and audio branding consultants are going to develop a new audio logo, they might reflect on how to use and combine musical parameters as efficiently as possible in accordance with the chosen strategy of building up brand awareness.

As for the ability to *recognise versus identify* the five brands, we have noticed that there is no one-to-one correspondence. Even though the number of students identifying or nearly identifying the brand approximates (with only, on average, a difference of five per cent in each test condition) the number of students ending up recognising the brand (cf. Fig. 3), there are notable *variances* among the individual logos (cf. Fig. 4). Whereas the level of brand identification corresponds completely with the level of brand recognition in the cases of McDonald’s and Nokia and almost (except for three per cent) in the case of Elgiganten, the difference is considerable in the case of JYSK and strongly marked in the case of Intel. Furthermore, and aside from these variances, the Intel and Elgiganten logos have been nearly identified *rather than* correctly identified, and that is different from the other three logos, which, by the way, happen to be the ones with pre-existing musical material. The latter is indeed interesting because it suggests that licensing or ‘remediating’ music might pay off, despite the fact that ‘a musical brand is typically both exclusive and unique – features, which pre-existing music lacks’ (Graakjær, 2009, pp. 108-109).

Considering the above-mentioned individual variances, it seems at first obvious that a high level of melodic distinctiveness is a conducive factor to the increase of logo and brand recognition as well as brand identification. However, since melodic distinctiveness is a multifaceted concept, involving at least five (and conceivably a few more) quantitative variables as well as a number of qualitative variables (cf. Table 1), it is a complicated matter to decide whether, for instance, the JYSK logo is more distinctive than the Elgiganten logo or vice versa. Add to this the unknown variables, such as the students’ previous acquaintance with the logos. Therefore, we

will restrict our conclusion to maintain that *the responses arguably suggest a correlation* between recognition on the basis of RAF and melodic distinctiveness.

Now, finally returning to the issue of *coding orientation*, it is evident that the concept of RAF not only implicates abstract reasoning, but also sensory processing, following Van Leeuwen (1999, pp. 177-178), since the concept allows for extensive 'emotion making' expression forms in terms of pitch range and durational variation (or rhythm) only. Still, it is a highly abstract kind of sonic and musical expression, which might appeal to only a limited number of subjects. When considering the habitual character of the students, it is important to take into account that they represent together a specialised sociocultural group, situated in a learning milieu, and as such there are perhaps reasons to assume that their perception of abstract stimuli is supported by the discursive context. Perhaps the students were more receptive than normal?

References

- Adlam, D.S. (1977). Introduction. In: Adlam, D.S., Turner, G., & Lineker, L. (Eds.), *Code in Context* (pp. 1-37). London: Routledge & Kegan Paul.
- Arning, C., & Gordon, A. (2006). Sonic Semiotics. The Role of Music in Marketing Communications. In: ESOMAR Annual Conference, London: *Congress 2006 – Foresight – The Predictive Power of Research* (pp. 1-18). London.
- Bang, K. (2011). *Lydlogoers genkendelighed. Lydbranding i et musikalsk perspektiv*. [The recognisability of sonic logos. A musical perspective on sonic branding]. Aalborg University. MA thesis.
- Bernstein, B. (1971). *Class, Code and Control: Volume 1. Theoretical Studies Towards A Sociology Of Language*. London: Routledge.
- Bernstein, B. (1981). Codes, Modalities, and the Process of Cultural Reproduction: A Model. *Language in Society*, 10(3), 327-363.
- Boeriis, M. S. (2009). *Multimodal Socialsemiotik & Levende Billeder* [Multimodal Social Semiotics and Moving Images]. University of Southern Denmark. PhD thesis.
- Bonde, A. (2013, forthcoming). Betydningen af melodi-interval og rytme for lydlogoers brand-tilknytning. In: Klingspor, E., & Derkert, J. (Eds.), *Proceedings of the 16th Nordic Musicological Congress 2012*. Stockholm.
- Bonde, A., & Bang, K. (2012). Musikalske lydlogoers DNA [The DNA of Musical Sonic Logos], *Akademisk Kvarter* [Academic Quarter].
- Bourdieu, P. (1977). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Bourdieu, P. (1986). *Distinction. A Social Critique of the Judgement of Taste*. London: Routledge.
- Bronner, K. (2004). Audio-Branding: Akustische Markenkommunikation als Strategie der Markenführung. Munich: Grin.
- Bronner, K. (2007). Schöner die Marken nie klingen... Jingle all the Way? Grundlagen des Audio-Branding. In: Bronner, K., & Hirt, R. (Eds.), *Audio-Branding: Entwicklung, Anwendung, Wirkung akustischer Identitäten in Werbung, Medien und Gesellschaft* (pp. 82-96). Munich: Reinhard Fischer.
- Bundsgaard, J. (2005). Bidrag til danskfagets it-didaktik. Med særligt henblik på kommunikative kompetencer og på metodiske forandringer af undervisningen. [Contribution to the Educational Theory and Practice of IT in the Danish Subject in primary and lower secondary school]. Copenhagen: School of Education. PhD thesis.

- Chandler, D. (2013). *Semiotics for Beginners*. Aberystwyth University. Accessed 4 April 2013.
- Constantinou, O. (2005). Multimodal Discourse Analysis: Media, Modes and Technologies. *Journal of Sociolinguistics*, 9(4), 602-618.
- Constantinou, O. (2005). Multimodal Discourse Analysis: Media, Modes and Technologies. *Journal of Sociolinguistics*, 9(4), 602-618.
- Cook, N. (1998). *Analysing Musical Multimedia*. Oxford: Oxford University Press.
- Eckkrammer, E.M. (2004). Drawing on the Theories of Inter-semiotic Layering to Analyse Multimodality in Medical Self-counselling Texts and Hypertexts. In: Ventola, E., Charles, C. & Kaltenbacher (Eds.), *Perspectives on Multimodality* (pp. 211-226). Amsterdam: John Benjamin.
- Eco, U. (2001). *Experiences in Translation*. Toronto: University of Toronto Press.
- Gorlée, D.L. (1997). Intercode Translation: Words and Music in Opera. *Target*, 9(2), 235-270.
- Graakjær, N.J. (2009). The JYSK Jingle. On the use of pre-existing music as a musical brand. In: Graakjær, N.J., & Jantzen, C. (Eds.), *Music in Advertising. Commercial Sounds in Media Communication and Other Settings* (pp. 99-119). Aalborg: Aalborg University Press.
- Graakjær, N.J. (2011). Om musik i tv-reklamer for McDonald's-kampagnen *i'm lovin'it* [On music in TV commercials for the McDonald's campaign *i'm lovin' it*]. *Akademisk Kvarter [Academic Quarter]*, 2, 250-262.
- Graakjær, N.J., & Jantzen, C. (2009). Mapping Research on Music in TV Commercials. In: Graakjær, N.J., & Jantzen, C. (Eds.), *Music in Advertising. Commercial Sounds in Media Communication and Other Settings* (pp. 13-52). Aalborg: Aalborg University Press.
- Groves, J. (2007a). Sound Branding – Strategische Entwicklung von Markenklang. In: Meyer, H. (Ed.), *Marken-Management 2008/2009. Jahrbuch für Strategie und Praxis der Markenführung* (pp. 125-148). Frankfurt o. t. Main: Deutscher Fachverlag.
- Groves, J. (2007b). A Short History of Sound Branding. In: Bronner, K., & Hirt, R. (Eds.), *Audio Branding: Entwicklung, Anwendung, Wirkung akustischer Identitäten in Werbung, Medien und Gesellschaft* (pp. 40-51). Munich: Reinhard Fischer.
- Halliday, M.A.K. (1978). *Language as Social Semiotic*. London: Edward Arnold.
- Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Haverkamp, M. (2008). *Synästhetisches Design: Kreative Produktentwicklung für alle Sinne*. Munich: Hanser.
- Heider, F., & Simmel, M. (1944). An Experimental Study of Apparent Behavior. *American Journal of Psychology*, 57(2), 243-259.
- Hjelmslev, L. (1961 [1943]). *Prolegomena to a Theory of Language*. Madison: University of Wisconsin Press.
- Hoek, L.H. (1995). La transposition intersémiotique pour une classification pragmatique. In: Hoek, L.H., & Meerhoff, K. (Eds.), *Rhétorique et Image Textes en homage à Á. Kibédi Varga* (pp. 65-80). Amsterdam: Rodopi.
- Hofer, J. (2009). *Akustische Signale als Werbemittel im Hörfunk*. University of Vienna. MA.
- Jackson, D.M. (2003). *Sonic Branding: An Introduction*. London: Palgrave Macmillan.
- Jackson, D.M. (2006). Orchestrating a Sound Strategy. *Brand Strategy*, 199, 38-40.
- Jakobson, R. (1959). On Linguistic Aspects of Translation. In: Bower, R. (Ed.), *On Translation* (pp. 232-239). Cambridge, MA: Harvard University Press.
- Jekosch, U. (2005). Assigning Meaning to Sounds – Semiotics in the Context of Product-Sound Design. In: Blauert, J. (Ed.), *Communication Acoustics* (pp. 211-226). Berlin: Springer.
- Kivy, P. (1990). *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca, NY: Cornell University Press.
- Kress, G., & Van Leeuwen, T. (1996 [2006]). *Reading Images: The Grammar of Visual Design*. London: Routledge.
- Krugmann, D. (2007). Integration akustischer Reize in die identitätsbasierte Markenführung. *LiM Arbeitspapiere*, 27, ed. C. Burmann. Universität Bremen.

- Kusatz, H. (2007). Acoustic Brand Management. Increasing the Acoustic Perception of Brands. 2nd International VDT Symposium. Sounddesign. Forum am Schlosspark, Ludwigsburg.
- Lauer, C. (2009). Contending with Terms: “Multimodal” and “Multimedia” in the Academic and Public Spheres. *Computers and Computation*, 26(4), 225-239.
- Levinson, J. (1984). Hybrid Art Forms. *Journal of Aesthetic Education*, 18(4), 5-13.
- Lindstrom, M. (2005). *Brand Sense: Build Powerful Brands through Touch, Taste, Smell, Sight, and Sound*. New York: Free Press.
- Machin, D. (2007). *Introduction to Multimodal Analysis*. London: Hodder Arnold.
- Machin, D. (2010). *Analysing Popular Music: Image, Sound and Text*. London: Sage.
- Machin, D. (2011). Towards a Social Semiotic Approach of the Analysis of Emotion in Sound and Music. *The Public Journal of Semiotics*, 3(2), 152-174.
- Middleton, R. (1990). *Studying Popular Music*. Milton Keynes: Open University Press.
- Mitchell, W.J.T. (2005). There Are No Visual Media. *Journal of Visual Culture*, 4(2), 257-266.
- Nerpin, S., Veit, R., & Heller, M. (2007). The Sound of Vattenfall: Ein Markenversprechen wird vertont. In: Bronner, K., & Hirt, R. (Eds.), *Audio Branding: Entwicklung, Anwendung, Wirkung akustischer Identitäten in Werbung, Medien und Gesellschaft* (pp. 266-278). Munich: Reinhard Fischer.
- Nölke, S.V. (2009). *Das 1x1 des Audio-Marketings: Der Navigator für Audio-Branding und Audio-Interface-Design*. Cologne: Comevis.
- O'Toole, M. (1994). *The Language of Displayed Art*. London: Leicester University Press.
- Palghat, V. (2009). *Hearing, Remembering, and Branding: Guidelines for Creating Sonic Logos*. University of Cincinnati. PhD thesis.
- Peirce, C.S. (1907[1998]). Pragmatism. In: Peirce Edition Project (Ed.), *The Essential Peirce. Selected Philosophical Writings*. Vol. 2 (1893-1913). Bloomington: Indiana University Press.
- Roth, S. (2005). *Akustische Reize als Instrument der Markenkommunikation*. Wiesbaden: Gabler.
- Ryan, T.A., & Schwartz, C.B. (1956). Speed of Perception as a Function of Mode of Representation. *American Journal of Psychology*, 69(1), 66-69.
- Stefani, G. (1987). A Theory of Musical Competence. *Semiotica*, 66, 7-22.
- Straka, M. (2007). *Audio-Branding im aktuellen Kontext der Marken-Kommunikation: Zur Struktur und Funktion der Elemente von 'Corporate Sound'*. Hamburg: Diplomica.
- Thibault, P.J. (1991). *Social Semiotics as Praxis: Text, Social Meaning Making, and Nabokov's Ada*. Minneapolis: University of Minnesota Press.
- Van Leeuwen, T. (1999). *Speech, Music, Sound*. London: Palgrave Macmillan.
- Westermann, C. (2008). Sound Branding and Corporate Voice: Strategic Brand Management Using Sound. In: Hempel, T. (Ed.), *Usability of Speech Dialog Systems. Listening to the Target Audience* (pp. 147-155). Berlin: Springer.
- Wingstedt, J., Brändström, S., & Berg, J. (2010). Narrative Music, Visuals and Meaning in Film. *Visual Communication*, 9(2), 193-210.
- Winter, M. (2011). *A Sonic Trilogy. Sound, Ostranenie and Social Consciousness*. University of Brighton. MA thesis.
- Zappavigna, M.S. (2005). Tacit Knowledge in Communities of Practice. In: Coakes, E., & Clarke, S. (Eds.), *Encyclopedia of Communities of Practice in Information and Knowledge Management* (pp. 508-513). Hershey, PA: Idea Group.

Notes

- 1 See e.g. Gorrée's discussion of words versus music in opera and other vocal genres (1997, pp. 237-240).
- 2 Cf. Bronner (2004, pp. 39-40), Groves (2007a, pp. 135-147; 2007b, pp. 45-47) and Hofer (2009, pp. 57-58).

- 3 For an outline of audio touch points, see, e.g. Jackson (2003, p. 6), Bronner (2004, p. 70; 2007, p. 89), Nerpin et al. (2007, p. 270), Groves (2007a, p. 127, 131), Krugmann (2007, pp. 33-35), Kusatz (2007, pp. 2-3), Westermann (2008, p. 154) and Nölke (2009, p. 79).
- 4 The empirical data of the experiment have been analysed previously (Bang, 2011; Bonde & Bang, 2012; Bonde, 2013). However, due to new appreciations of how the experimental procedure might have an impact on the results, we offer a different interpretation in this paper.
- 5 Recently, there have been a few other contributions merging Hallidayan systemic-functional grammar and social semiotics into the field of music and sound, such as Wingstedt et al. (2010) and Machin (2010, 2011).
- 6 Additionally, when comparing the visual and the sonic modality cues (Fig. 1 and 2), one may notice prominent cross-sensory analogies such as *brightness* and *pitch range* or *depth* and *perspectival depth*.
- 7 Cf. Boeriis (2009, pp. 27-30) for a manageable outline of text and context (based on systemic-functional grammar), including the hierarchical layers: 'expression' (or 'sentence'), 'text', 'co-text', 'situation' and 'culture'. For a music semiotic perspective on 'coding' and 'decoding' as well as 'coding competence', see Stefani (1987) and Middleton (1990, p. 173f).
- 8 It appears implicitly in Van Leeuwen (1999) that technological coding orientation does not have a counterpart in the sound domain, and that seems surprising given that it is a well-known fact that there are plenty of examples of sounds with low articulatory complexity, which are meant for taking action: e.g. the sound of an alarm clock or the 'beep' sounds at traffic lights helping blind people distinguish between red light and green light when crossing the street.
- 9 Actually, even without discussing the emerging field of audio branding, Van Leeuwen implies such a link himself by proposing student exercises in advertising voice analysis and audio logo design (1999, p. 155).
- 10 Cf. Graakjær (2011) for a description of the genealogy of the McDonald's *i'm lovin' it* logo.
- 11 In the meantime, however, the Elgiganten logo has changed.
- 12 The audio files are linked to Example 3-7. By clicking with the mouse on the score sample, six different versions are sounded, one after another, in the following order: 1) pitches only, slow tempo; 2) rhythm only, half tempo; 3) pitches only, fast tempo; 4) rhythm only, original tempo; 5) pitch and rhythm, original tempo (RAF); and 6) original sound logo
- 13 As a starting point, the designations 'slow' and 'fast' relate to *longest* and *shortest* time distance (duration), respectively, between two pitches in the logo in question. However, seeing that some of the logos (in versions I and II) turn out to be undesirably fast or slow, the tempo differences are adjusted intuitively.
- 14 In Test 2 virtually nobody has recognised any of the five logos and brands on the basis of the first version. The only (and insignificant) exception is the McDonald's logo with a recognition rate of one per cent.
- 15 See Jackson (2006) who reports a similar divergence between audio logo recognition and brand identification.
- 16 The beginning minim and the succeeding four quavers in the Intel logo (cf. Example 3) could be defined, after all, as *two parts*. Though, with only one note sounding, the first part is hardly to be perceived as a structural unit itself.
- 17 Just as McDonald's has made the hook line of the Timberlake song their sounding signature, the 'Nokia tune' makes up a four-bar finishing phrase of Francisco Tárrega's 'Gran Vals' for guitar solo (1902). As for the JYSK logo, the seven-note phrase originates from an old Danish patriotic song, 'Man from Jutland' (1846), in which the sung text, 'The Man from Jutland is strong and tough', is associated with austerity and unpretentiousness personified by the owner of the company, Lars Larsen (cf. Graakjær, 2009, p. 100f).

Example 3



Example 4



Example 5



Example 6



Example 7

