LURE(D) INTO LISTENING: THE POTENTIAL OF COGNITION-BASED MUSIC INFORMATION RETRIEVAL

Henkjan Honing
Institute for Logic, Language & Computation (ILLC)
Cognitive Science Center Amsterdam (CSCA)
University of Amsterdam
www.hum.uva.nl/mmm
honing@uva.nl

ABSTRACT

This paper argues for the potential of cognition-based music retrieval by introducing the notion of a musical 'hook' as a key memorization, recall, and search mechanism.

1. INTRODUCTION

A common approach in Music Information Retrieval (MIR) is to use information-theoretic models to extract information from the data. It is based on the assumption that all relevant information is present in the data itself and that it can, in principle, be extracted (*data-oriented approach*). However, with regard to the cognitive aspects of music (the perspective of the listener), some information might not be explicit or present at all in the data. For example, two silences in a fragment of music might be physically identical, while for a listener one silence might be perceived as highly salient, because it is in a position that is metrically important (a so-called 'loud rest') as compared to a non-salient silence in a metrically unimportant position [8]. Elaborating state-of-the-art MIR techniques with recent findings from music cognition seems therefore a natural next step in improving (exploratory) search engines for music and audio (*cognition-based approach*).

This paper will focus on the most salient, memorable, and easy to recall moment of a musical phrase or song, the so-called 'hook', as a way to understand and identify which cognitively relevant musical features affect the appreciation, memorization and recall of music. To illustrate the potential of this idea, in the second half of the paper a pilot project is described.

2. THE 'HOOK' OF A MELODY

The idea to focus on the 'hook' in music in order to improve the state-of-the-art automatic music annotation and MIR techniques is novel. While the term is common among musicians and musicologists (cf. [3]), what precisely constitutes a hook and what makes it stick in our minds is unknown [10]. While plenty of empirical data is available from mostly questionnaire-style research (cf. [2]), how in fact the musical structure —the moment in time that 'carries' the hook— is, as yet, unclear. Hence there is no theory to build on. Therefore in this paper we propose a novel, yet relatively straightforward annotation technique in which a large quantity of listeners mark and annotate music with what they consider the 'hook', the essential or catchy part of a song. Given the audio recording, information about the musical structure (e.g., pitch and rhythm), and the listeners descriptions of the particular moment in time that stood out —emotionally, structurally, perceptually, etc.—, it becomes possible to characterize what actually establishes a hook.

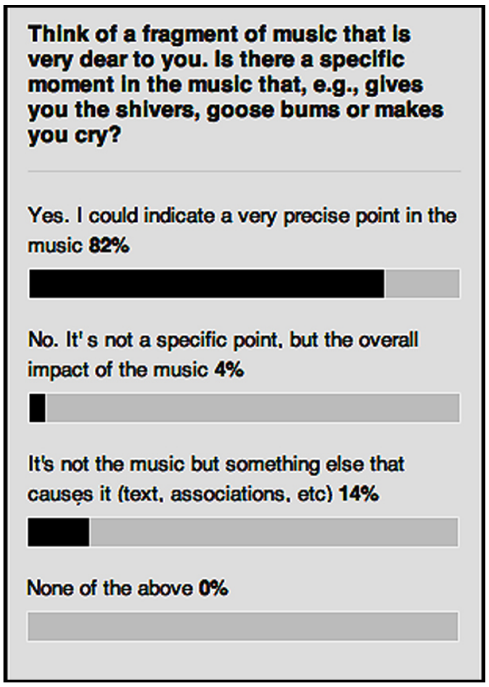

Figure 1. Result of an informal Web poll that was online during the preparation of this paper [16].

The availability of a model of a hook will help in addressing a key problem for musicologists: trying to relate transcription and/or recordings of songs that have been transformed over the years. This becomes relevant when,

for instance, one tries to establish the interrelationships of composed popular music and orally-transmitted folk songs (cf. [6]). For both aspects ones needs to understand which cognitively relevant musical features affect the appreciation and memorization of music.

To realize this, one has to identify, first, which features might be relevant for listeners with respect to appreciation, memorization and recall of that particular musical moment. Second, one needs large amounts of semantically annotated musical data. We propose to obtain such data by collecting information from large numbers of listeners via a web-based environment (the *locator*) where listeners are encouraged to mark specific locations in an actual recording, locations where s/he experienced something special or that s/he considers musically striking or intriguing (see Section 3.4 for a more elaborate description).

In addition, large quantities of annotated musical data will allow for evaluating the explanatory power of cognitive models of melody and rhythm perception, insights about the way people remember a melody and recall such a melody from memory (e.g., [1], [10]). The availability of these empirical data has great potential in revealing what structural (e.g., pitch, key, rhythm, meter) and non-structural (e.g., perceptual, associative, emotional, cultural) aspects of a melody play a role in the memorization and experiencing of music.

Based on the ideas outlined above, below an example is given of an interdisciplinary research project where scholars and scientists from musicology, media studies, psychology and computer science plan to collaborate. It might serve as a concrete example of the potential of this approach in improving MIR search engines and inform music cognition theories of the perception and memorization of music.

3. PILOT PROJECT: *LURELEI*

3.1 Overview: *LISTEN, LURE & LOCATE*

This research project (in preparation [17]) is centred around the idea about being lured into, or luring others to listen to music. It aims to study the phenomenon of being lured to listen to new, unfamiliar musics, and especially the role that internet-mediated technologies can have in this process. In short, the research aims to analyze and explicate, to design and construct, as well as to observe and evaluate truly engaging, yet virtual meeting points for music listeners. Such a virtual listening space (VLS for short) will allow participants to share their listening experiences (*LISTEN*), make other listeners enthusiastic for a certain musical fragment (*LURE*), and mark a specific location in an actual recording (*LOCATE*) — a specific point in the music where a particular listener experienced something special or that s/he considers musically striking or intriguing (cf. [13]).

3.2 *LISTEN*: The activity of interest

This part of the project will focus on the act of music listening, the diversity and change of listening modes (from concert hall to headphone), and how they are influenced by the availability of recording and playback techniques (from phonograph to iPod). With the arrival of internet ('Web 2.0') specific listening communities and listening cultures have emerged revealing their own kind of dynamics. The internet provides a novel opportunity to trace listening experiences, their communication, and how these relate to the actual content of the music itself. The challenge is to bridge recent insights from cultural, cognitive and media studies [5, 9, 11, 14] to the field of MIR in which these VLSs can be built, analysed and empirically studied.

3.3 *LURE*: The effect of engaging 'criticism' by experienced listeners

The second part of the research project investigates older and newer internet technologies that are concerned with user-generated content, environments that aim at sharing musical taste and exchange of musical listening experiences. The consortium is especially interested in how existing internet communities (e.g., *Last.fm*, *Pandora*, *YouTube*, *Spotify*) have an effect on musical preference as compared to traditional 'music criticism' present in the more traditional media like newspapers, radio and television [4]. Next to the apparent social and cultural function of these environments, the focus will primarily be on the listening experience that is shared in these virtual meeting points, by looking at how specific listening experiences are communicated among users, and how both the experiences and their communication relate to the actual content of the music itself. This study will address research questions including: Does the direct sharing of personal listening experiences amount to a new kind of music criticism? Which concepts and criteria allow for evaluating this new type of popular criticism (cf. the phenomenon of blogs)? What makes these virtual listening spaces more compelling —luring listeners into new modes and domains of musical listening— than other media? How do people find each other on the internet, and why do they choose a particular website? What is the role of the technology in steering this community-sharing and community-building?

3.4 *LOCATE*: Relating the listening experience to musical structure

The third part of the research project aims to actively experiment with 'Web 2.0' technologies by designing and constructing a VLS that will support the sharing of listening experiences.

The *LOCATE*-component of the VLS is actually a novel idea. It allows the listener to pinpoint to the specific moment in the music that s/he considers special, captivat-

This is research in progress (January 2010). An updated paper can be found at: http://www.hum.uva.nl/mmm/abstracts/honing-2010a.html

ing, etc. This focus on a specific location in the music that functions as a 'hook' (both musically and cognitively) to a description of a strong listening experience is the basis of the VLS. This will allow for evaluating —in an empirical and controlled way— the potential of 'Web 2.0' in sharing listening experiences (cf. [7]), and it enables us to advise on how these might be used in the creative industry.

The ambition is to design and develop a novel infrastructure to let listeners collectively provide annotations and to derive cognitively relevant features (the *locator*). The system will allow for the following research questions to be addressed: (i) which are the candidate 'hooks' of a melody, i.e. the fragment of the melody most people remember, or will start singing when asked to do so. As such a 'hook' can be considered the 'essence' of a song, and might facilitate search in a large database of songs. (ii) How can existing measures of melodic similarity be enhanced with a measure of rhythmic similarity? Since a 'hook' is composed of melodic as well as rhythmic information, both need to be captured in a model to be able to use it for annotation and search.

On the bases of these derived, yet cognitively relevant features (the 'hook', melodic and rhythmic similarity, etc.) content-based retrieval methods for musical audio will be developed. Outstanding research questions are (i) how to map the cognitively relevant features into computational models, (ii) how to design and evaluate similarity measures for these features, and (iii) to select efficient and robust algorithms that allow to compute them.

3.5 *LISTEN*, *LURE* & *LOCATE*: A meeting point

In addition to the core projects mentioned above, the project allows for more general questions of interest to the *computational humanities* [15]. For instance, what are the consequences of internet, both in terms of transmission and dissemination as well as the promotion of listening experiences, for music cognition and appreciation? What is the influence of the new emerging forms of criticism on the perception of musical nuance? And in what way do listening experiences, that are mediated through the internet, differ from 'real' listening spaces like public spaces or concert halls? And last but not least, the empirical data obtained forms a solid starting point for *cognitive science* in studying what could explain that some melodies behave like 'earworms' and others don't. That is, what structural aspects make these melodies spontaneously appear and consequently stick in one's mind? [18, 10]

4. CONCLUSION

This paper argued that the combination of *crowd annotation* (i.e., social- or crowd-tagging) and marking the specific moment in ones favourite music that carries a strong emotion, the *hook*, has great potential for improving search engines for music. In addition the annotations will provide a rich empirical source to music cognition research in underpinning what makes certain melodic fragments more sticky than others.

5. ACKNOWLEDGEMENTS

Parts of this paper were used in an earlier grant proposal that benefited from comments from several researchers, including José van Dijck, Remco Veltkamp, Frans Wiering, Louis Grijp, and Simon Frith.

6. REFERENCES

[1] Baddeley, A.D. (2000) Short-term and working memory. In: E. Tulving & F.I.M. Craik (Eds) *The Oxford Handbook of Memory*, New York: Oxford University Press, 77-92.

[2] Bennett, S. (2002). *Musical Imagery Repetition (MIR).* Master thesis, Cambridge University.

[3] Burns, G. (1987). A typology of 'hooks' in popular music. *Popular music,* 6(1), 1-20.

[4] Frith, S. (2003). Music and everyday life. In Clayton, M. et al. (ed.) *The Cultural Study of Mu*sic, 92-101. London: Routledge.

[5] Frith, S. (2008). Why music matters, *Critical Quarterly,* 50, 165-179.

[6] Grijp, L.P. (2008). Onder de altijd groene linde. Over orale principes in Middelnederlandse liederen. In: L.P. Grijp & F. Willaert (Eds.), *De fiere nachtegaal. Het Nederlandse lied in de middeleeuwen*. Amsterdam University Press, Amsterdam. 311-329.

[7] Honing, H. (2006). On the growing role of observation, formalization and experimental method in musicology. *Empirical Musicological Review*, 1(1), 2-5.

[8] Honing, H. *et al.* (2009). Is beat induction innate or learned? Probing emergent meter perception in adults and newborns using event-related brain potentials (ERP). *The Neurosciences and Music III — Disorders and Plasticity: Annals of the New York Academy of Sciences*, 1169, 93–96.

[9] Honing, H., & Reips, U.-D. (2008). Web-based versus lab-based studies: a response to Kendall (2008). *Empirical Musicology Review*, 3 (2), 73-77.

[10] Hubbard, T. L. (2010). Auditory imagery: empirical findings. *Psychological Bulletin*, *136*(2), 302-329.

[11] Juslin, P., & Sloboda, J. (Eds.) (2009). *Handbook of Music and Emotion: Theory, Research, Applications.* Oxford: University Press.

[12] Levitin, D.J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Perception & Psychophysics* 56, 414-423.

[13] Sloboda, J.A. (1991). Music structure and emotional response: some empirical findings. *Psychology of Music*, 19, 110–120.

[14] Van Dijck, J. (2006). Record and hold: Popular music between personal and collective memory. *Critical Studies in Media Communication*, 23 (5), 357-374.

[15] Willekens, F. *et al.* (in preparation) *Computational Humanities. Roadmap to the Humanities in 2025*. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschappen (KNAW).

[16] http://www.musiccognition.nl/blog/

[17] http://www.musiccognition.nl/lll/

[18] http://www.hum.uva.nl/mmm/oorwurm/ (in dutch)

This is research in progress (January 2010). An updated paper can be found at: http://www.hum.uva.nl/mmm/abstracts/honing-2010a.html