UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Lure(d) into listening: The potential of cognition-based music information retrieval.

Honing, H.

Publication date
2010
Document Version
Final published version
Published in
Empirical Musicology Review
License
CC BY-NC

Link to publication

Citation for published version (APA):
Honing, H. (2010). Lure(d) into listening: The potential of cognition-based music information retrieval. *Empirical Musicology Review*, *5*(4), 121-126.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Sep 2026

Lure(d) into listening: The potential of cognition-based music information retrieval

Henkjan Honing
University of Amsterdam

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. A hook is considered the most salient, memorable, and easy to recall moment of a musical phrase or song. Next to its role in searching large data-bases of music, it is proposed 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 for the *computational humanities* (Willekens et al., 2010), in the second half of the paper a pilot research project is described. This project, named *Listen, Lure* & *Locate*, aims to study the cultural phenomenon of being lured to listen to new unfamiliar music, and especially the role that recent internet-mediated technologies can have in this process. It is argued that a combination of *crowd annotation* (i.e., social- or crowd-tagging) and marking the specific moment (the hook) in one's favorite music, has great potential for improving search engines for music. In addition, these annotations will provide a rich empirical source to music cognition research in determining what makes certain melodic fragments more sticky than others.

Submitted 2010 May 28; accepted 2010 July 16

KEYWORDS: *music cognition, music information retrieval, computational humanities, hook, crowd annotation*

INTRODUCTION

The widespread availability of the Internet and the development efficient compression techniques like MP3 audio encoding in the last ten years, gave an enormous boost to a discipline that is now referred to as Music Information Retrieval (Downie, 2003). In that discipline the topic is not so much to understand and model music, but to design robust and effective methods to locate and retrieve musical information, including tasks like query-by-humming, music recommendation, music recognition, and genre classification.

A common approach in Music Information Retrieval (MIR) is to use information-theoretic models to extract information from the musical data, be it the audio recording itself or all kinds of meta-data, such as artist or genre classification. With advanced machine learning techniques, and the availability of so-called 'ground truth' data (i.e., annotations made by experts that the algorithm uses to decide on the relevance of the results for a certain query), a model of retrieving relevant musical information is constructed. Overall, this approach is based on the assumption that all relevant information is present in the data and that it can, in principle, be extracted from that data (*data-oriented approach*).

Several alternatives have been proposed, such as models based on perception-based signal processing (Aucouturier & Pachet, 2002) or mimetic and gesture-based queries (Lesaffre et al., 2004). However, with regard to the cognitive aspects of music information retrieval (the perspective of the listener), some information might be implicit or not 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'; London, 1993) as compared to a non-salient silence in a metrically unimportant position. Especially in the design of similarity measures (e.g., 'search for songs that sound like X') it becomes quickly clear that not all required information is present in the data (Aucouturier & Pachet, 2002; Lesaffre et al., 2004). 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 proposes to develop a cognition-based search engine based on the most salient, memorable, and easy to recall moment of a musical phrase or song, the so-called 'hook', in an attempt to identify which cognitively relevant musical features affect the appreciation, memorization and recall of music. To illustrate the potential of this idea for the computational humanities (Willekens et al., 2010), media studies and music cognition, in the second half of the paper a pilot research project is described.

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. Burns, 1987), what precisely constitutes a hook and what makes it stick in our minds is unknown (Hubbard, 2010). While plenty of empirical data is available from mostly questionnaire-style research (Bennett, 2002), What role the musical structure plays in this, i.e. the moment in time that 'carries' the hook, is still 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 to be 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.

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 (e.g., Grijp, 2008). 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 *Locate*-component, see below) 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 below 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., Baddeley, 2000; Hubbard, 2010). 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, an example is given below 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.

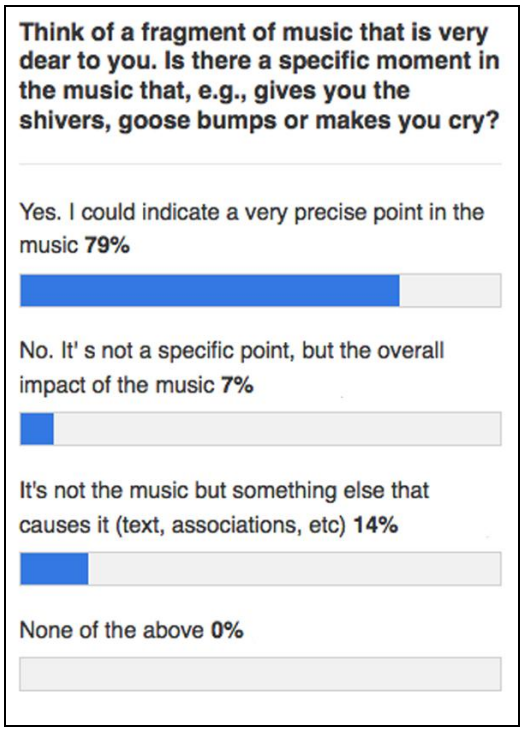

Fig. 1. Result of an informal Web poll (N=148) that was online during the preparation of this paper. [2].

EXAMPLE PROJECT

Overview: *Listen, Lure & Locate*

This research project (in preparation [1]) plans to investigate how people are being lured, or lure others to listen to music. It aims to study the phenomenon of being lured to listen to new and unfamiliar musics, focusing on 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 about 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. Sloboda (1991) showed that most music lovers actually have such experiences. They can often indicate the precise location in the music that 'carries' their experience or memory (See Figure 1 for the results of an informal poll on this topic).

Listen: The activity of interest

This part of the project aims to 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 (e.g., Frith, 2008; Honing & Reips, 2009; Juslin & Sloboda, 2009; Van Dijck, 2006) to the field of MIR in which these VLSs can be built, analyzed and empirically studied.

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*, *You-Tube*, *Spotify*) have an effect on musical preference as compared to traditional 'music criticism' present in the more traditional media like newspapers, radio and television (Frith, 2003). 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?

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 a novel idea. It allows the listener to pinpoint the specific moment in the music that s/he considers special, captivating, 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 the evaluation —in an empirical and controlled way— of the potential of 'Web 2.0' in sharing listening experiences (cf. Honing, 2006), 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 that permits listeners collectively to provide annotations and to derive cognitively relevant features (the *locator*). The system will allow the following research questions to be addressed: (i) What 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.

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. For instance, what are the consequences of the 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 will form a solid starting point for cognitive science in studying what could explain that some melodies behave like 'earworms' and others do not. That is, what structural aspects make these melodies spontaneously appear and consequently stick in one's mind? [3]

CONCLUSION

We argued that the combination of crowd annotation (i.e., social- or crowd-tagging), and marking the specific moment in one's favorite music that has a high level of 'stickiness', the hook, has great potential for improving search engines for music. In addition, the described methodology will provide a rich empirical source to music cognition research in determining what makes certain melodic fragments more

sticky than others. And finally, the example project might serve as a pilot research program where very different disciplines (e.g., media studies, music cognition and computational humanities) collaborate to further understand the social, cultural, cognitive and internet-mediated processes that are involved in listening to music.

ACKNOWLEDGEMENTS

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

NOTES

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

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

[3] http://www.hum.uva.nl/mmm/oorwurm/ (in Dutch)

REFERENCES

Aucouturier, J.-J., & Pachet, F. (2002). Finding songs that sound the same. *Proceedings of the IEEE Benelux Workshop on Model-Based Processing and Coding of Audio* (pp. 91–98). Leuven, Belgium: University of Leuven.

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, pp. 77-92.

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

Burns, G. (1987). A typology of 'hooks' in popular music. *Popular Music,* Vol. 6, No. 1, pp. 1-20.

Downie, J.S. (2003). Music Information Retrieval. *Annual Review of Information Science and Technology*, Vol. 37, pp. 295-340.

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

Frith, S. (2008). Why music matters, *Critical Quarterly,* Vol. 50, pp. 165-179.

Grijp, L.P. (2008). Introduction. In L.P. Grijp & I. van Beersum (Eds.), *Under the Green Linden — 163 Dutch Ballads from the Oral Tradition* (pp. 18–27). Amsterdam: Meertens Institute.

Honing, H. (2006). On the growing role of observation, formalization and experimental method in musicology. *Empirical Musicological Review*, Vol. 1, No. 1, pp. 2-5.

Honing, H., & Reips, U.-D. (2008). Web-based versus lab-based studies: a response to Kendall (2008). *Empirical Musicology Review*, Vol. 3, No. 2, pp. 73-77.

Hubbard, T.L. (2010). Auditory imagery: empirical findings. *Psychological Bulletin*, Vol. 136, No. 2, pp. 302-329.

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

Lesaffre, M., Leman, M., De Baets, B., & Jean-Pierre Martens, J-P. (2004). Methodological considerations concerning manual annotation of musical audio in function of algorithm development. *Proceedings of the 5th International Conference on Music Information Retrieval* (pp. 64-71). Barcelona, Spain: Pompeu Fabra University.

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

London, J.M. (1993). Loud rests and other strange metric phenomena (or, meter as heard). *Music Theory Online*, Vol. 0, No. 2. http://mto.societymusictheory.org/issues/mto.93.0.2/mto.93.0.2.london.art.

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

Van Dijck, J. (2006). Record and hold: Popular music between personal and collective memory. *Critical Studies in Media Communication*, Vol. 23, No. 5, pp. 357-374.

Willekens, F. *et al.* (2010). *Computational Humanities. Roadmap to the Humanities in 2025*. Amsterdam: Royal Netherlands Academy for Arts and Sciences (KNAW).