A TURING TEST FOR THE SINGING VOICE AS AN ANTHROPOLOGICAL TOOL: EPISTEMOLOGICAL AND TECHNICAL ISSUES

George Kosteletos
National and Kapodistrian
University of Athens
School of Philosophy
Department of Philosophy,
Pedagogy and Psychology
Athens, Greece
georgekosteletos@hotmail.com

Anastasia Georgaki
National and Kapodistrian
University of Athens
School of Philosophy
Department of Music Studies,
Athens, Greece
georgaki@music.uoa.gr

ABSTRACT

In this paper we propose the design and implementation of a Turing Test (TT) for the research of the singing voice. Although the TT is mainly related to the research field of Artificial Intelligence (AI), being used both as a criterion and an operational guide by the scientists of this field, with the present paper we attempt to introduce a rather different approach to the TT. Given the fact of various disputes over the validity of the TT as a criterion of AI, one might argue that the TT is nothing more than a 'philosophical fossil', a left-over and remainder of past and outdated philosophical assumptions about the nature of human intelligence. The problem of an unavoidable subjectivity in the results of TT experiments has strengthen the question about the usefulness of the TT as a research means. Our goal is to introduce a new scope for the use of the TT not as a criterion of intelligence but as an 'instrument' for tracing certain features of human judgment in various fields. Pretty much in the fashion of a Transcendental philosophical stance, we face the TT as a procedure in which what is judged is judgement itself. Specifically, in the present paper we attempt to exhibit the way in which a TT can be used to trace and highlight features of human judgment regarding the singing voice. Are certain factors like culture, gender, age or familiarity with music technology basic parameters of the way in which humans perceive and judge artificial and natural singing voice? Is the TT a worn off chapter in the history of the philosophy of AI or could it be a brand new tool for the research in fields like psychoacoustics, cognitive musicology, social psychology of music and generally for the research on research itself?[1] This is the kind of questions that we intend to raise concerning the future of the TT, starting with a paper for a possible 'singing voice TT'.

1. INTRODUCTION

1.1. Turing Test and Artificial Intelligence

The TT is named after the famous mathematician Alan Turing, since it was introduced by him in an article titled '*Computing, Machinery and Intelligence*' which meant to be a milestone for the field of AI. Opening his article with the intriguing question 'Can machines think?'[103] Turing went on to examine the possibility of tracing intelligence in mechanic entities (and finally in any entity) via written language communication.

Since then, '*Computing, Machinery and Intelligence*' has become one of the most referenced, republished and paraphrased philosophical articles that have ever been produced [41], [87], [23] and has actually prepared the ground for the 'Physical Symbol System Hypothesis' [94], [72] which became the basic Paradigm[2] for the 'symbolic AI' research program [34].

Since 'Symbolic AI' started to retreat-during the late 70s and early 80s-and AI researchers have progressively begun to turn to the 'connectionist approach' and generally to more holistic ways of viewing intelligence [33], [34], appropriate modifications have also been proposed for the TT [31], [43], [44], [68], [101].

In the completely opposite direction have moved researchers like Ray Kurtzweil who proposes the specialisation of the TT in very specific domains of intelligent action [60], [5].

1.2. The Musical Turing Test

Sooner or later, the designers of any version of TT have to deal with the question: On what exact measure should our TT be based? In other words, on which grounds

[1] If the TT was to unravel the basic features of human judgment, it could become an important 'vehicle' of self-knowledge and self-understanding for practically any researcher.

[2] We use the word "Paradigm" with the meaning that is given to it be Thomas Kuhn in his '*The Structure of Scientific Revolutions*', that is to refer to a theory which is used as a prototype, as a basis for the design and implementation of experiments . In order to distinguish from its usual meaning we've written the word with a capital "P".

should we evaluate what is observed by us as outcome of our TT? How should we filter these observations? *Originality* of action and finally *creativity* are nowadays very popular measures but they have been proposed as measures from the very first days of the TT and even before that.

In the history of Cognitive Science originality and creativity are mentioned for the first time by Lady Lovelace. Commenting on the limits of Charles Babbage's 'Analytic Engine', Lovelace expressed the view that a truly intelligent machine should present original-therefore unexpected-action [64]. Lovelace's view is commented by Turing in '*Computing, Machinery and Intelligence*' [103] and it is in this article that we find another famous early view regarding creativity and originality: "Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by false chance of symbols, could we agree that machine equals brain" [56], [103].

Artistic creativity seems also to satisfy those who demand a holistic approach to the problem of evaluating an observed action in terms of intelligence [15], [38], [50], [60].

From all the forms of art, music seems to be the one that mostly attracts the interest of the TT designers. After all, as mentioned by Gareth Loy, the abstract nature of music, therefore its susceptibility to mathematics and finally to computing, is the feature that makes music appropriate medium for a TT [65]. Belgum was the first who proposed the adoption of music as appropriate medium for the TT [12], while Pearce, Meredith and Wiggins report that one of the basic motives for the production and use of generative music systems is the development of 'tools' for the verification of certain theories of Cognitive Science [79].

The interest of the AI community for the use of music as a medium for the TT has led to a variety of 'musical TTs'[1]: "Musical Directive toy Tests" (MDtT) [5], [26], [28], [81], "Musical Output toy Tests" (MOtT) [11], [13], [40], [47], [61], [66], [69], [74], [93], [97], 'stylistic MDtT's' [75], [78], [117], 'musical Total TTs' (musical TTT), the musical version of Harnad's 'TotalTT'[12], [98], [109].

What use could all this research have?

2. THE TURING TEST: FROM A CRITERION TO A COMMENT. FROM A COMMENT TO A RESEARCH 'INSTRUMENT'

Apart from those who propose, design and organise TT's, there are also researchers who doubt the efficiency of the TT as a criterion for intelligence.

In our view, the critique against the TT can be divided into four major problems. First there is the problem of what we could call 'copyright of intelligence' or in other words 'authorship'. When the critics in a TT observe an action that they characterise as intelligent, to whom should they attribute the origins of this action? To the observed computing system or to its programmers? There is always the possibility of 'mimesis' [115] and in this case, what is intelligent is not the observed system but its programmers [41]. Ariza thinks of the TT as insufficient due to the fact that it encloses the idea of deception [5]. Even some researchers involved in the TT research deny that deceiving the critics has any relation with the potential of a computer to present intelligent behaviour [81], [65], [66], [26], [27], [28], [41], [86]. This problem of deception regarding the authority was already mentioned by Turing himself [103] and can be related to what Searle attempts to state with his 'Chinese Room Argument': pure *syntax* can produce the same behavioural result with *meaning* [88]. Thus syntax can deceive us and pass as being meaning (therefore intelligent treatment of symbols) [41], [30], [108]. As Boden states, the success of a machine in a TT is rather a matter of our own political and ethical decision [14].

The latter statement brings us to the second problem of the TT, which is the unavoidable subjectivity in the judgment of the critics. In a Wittgensteinian manner we would say that the critics participate to a certain "form of life" therefore their judgment is guided by the values of this "form of life". On can find a similar line of critique in the work of French [38], while Michie stresses the importance of "social intelligence" and proposes a 'TT for social intelligence' [68]. Again, Turing had already foreseen this problem when he mentioned that surprise-therefore alignment or not with our social stereotypes-can easily be considered as originality [103]. The effects of cultural subjectivity were shown clearly in the mistaken judgment reported by Halpern [41] and Churchland [19] considering a Loebner Test that took place in 1991.

The subjectivity of human judgment is doubled when we adopt art-for instance music-as a medium of our TT. This is actually the third problem regarding the TT which is specified as a problem of all the 'artistic TTs': an aesthetical subjectivity which presents its self both as a subjectivity in the notion of art and as a subjectivity of the artistic taste. Daligh and Schaffrath have actually organised a series of TT's the results of which showed that critics' aesthetical preferences affected their judgment regarding the intelligence of the participants [29]. Same work has been carried out by Soldier [97]. Wiggins observes that in a 'musical TT' the definition and tracing of intelligence results in a problem equal to the problem of defining "good music": a problem of aesthetics [112]. Similar views are expressed by Laurie Spiegel [12] and David Cope [28].

But embracing art as a medium for the TT can reveal another problem which is general to all versions of the TT. Back in the 16th century Shaftesbury observed that there can be aesthetical objects that are not artistic objects, they are not made by humans at all and are simply physical objects made by nature [10]. Shaftesbury's observation points at something that in

[1] An overview of all the proposed 'musical TTs' as well as all the proposed TT in general is presented by Ariza [5].

modern philosophy is marked out by Dennett as "intentional stance": humans' tendency to personalize everything, therefore to attribute intentions (artistic, ethical etc) to entities that are considered 'soulless' like machines are [32], [51]. Hofstadter has called our intentional stance towards the AI computing systems as "ELIZA effect" [49], while Ariza points out that intentional stance could make the critics of a 'musical TT' attribute intelligence to non intelligent entities [5].

The above analysis of the problems of the TT has turned the focus of our discussion from the evaluated participants to the evaluating interrogator. Our interest seems to be shifted from the ontology of the evaluated entity to the ontology of our evaluation process itself. In our view this shift can be already found in the proposals for 'Musical Discrimination Tests' (Musical DT's) [29], and 'reverse TTs' (or 'fully automated TTs' or 'CAPTCHA') [71], [6], [106], [107]. In this sense, the AIBO project, the 'embodied TT' designed and executed by Swisher, Dotov and Chemero [101] is also a TT that focuses on the interrogator's judgment since its goal was the statistical analysis of the attribution of ethical responsibility to a robot by human interrogators.

This turn of interest from the participants to the interrogators might do justice to Turing's initial intentions. Some researchers believe that with '*Computing, Machinery and Intelligence*' Turing did not intend to introduce a sufficient criterion of intelligence but rather wanted to trigger a discussion about the subjectivity of human judgment regarding the attribution of intelligence to others [45], [68], [59], [111] or even about the subjectivity of human judgment in general [46]. After all, Turing himself stated that by the end of the 20th century at least the educated people will have changed their views regarding the definition of intelligence and the potential for machine intelligence, not because of a technological progress but because of a change in our beliefs [103]. Thus, Turing's introduction of a TT does not seem to be a 'manual guide' for the mechanical reproduction of intelligence, but is rather an exploration of the mechanical potential for the satisfaction of our stereotypes, finally a statement for the arbitrariness of human judgment, a critique of human judgment.

We believe that this rather Kantian way of thinking of the TT is the only way in which the TT can be proved fruitful in any research, that is as a means of revealing the way our subjectivity is structured and employed so, in other words as an 'anthropological tool'. This is exactly the treatment of the TT that we intend to adopt in our proposal for a 'singing voice TT'.

3. A TT FOR THE SINGING VOICE

3.1. Theoretical background

It is striking that, in the above mentioned research on TT, none TT for singing voice has been carried out until now. There have been only some TTs regarding speech [20], [106], [63], [18], [57], [91].

We propose the design and implementation of a 'singing voice TT' as an extension and follow-up work of the psychoacoustic experiment carried out by the singing voice research group of Kouroupetroglou and Georgaki, for the evaluation of a Greek voiced score-to-singing voice synthesis system [62]. In that experiment, 30 human participants, aged between 20 and 50 years old, evaluated and compared an older Greek diphone database (GR2) produced by using the MBROLA synthesizer with a new improved and extended Greek diphone database (GR3) produced also with MBROLA. The experiment was actually a test of MOS (Mean Opinion Score) in which participants evaluated two singing voice samples based on GR2 and GR3 respectively according to signing voice 'qualities' characterised as "vocalness", "naturalness", "intelligibility" and "expressivity".

Nevertheless, that experiment lacked for a comparison between synthetic and natural singing voice, since both the evaluated samples were synthetic. GR3 was found better than the GR2 in all aspects but a complete evaluation of GR3 demands its comparison with samples of natural singing voice. It is exactly the comparison between mechanic (synthetic) and human (natural) samples that could give to our experiment the 'flavour' of a TT.

But what could be really evaluated with this TT? First of all, this is a test that does not aim at the evaluation of intelligence but seems to aim at the evaluation of aesthetical quality. Therefore it already presents a significant difference to what is usually believed to be the original conception of the TT. On the other hand, as we presented in section 2, it seems that Turing's intentions concerned not the ontology of the evaluated entities but the ontology of the evaluation itself. Therefore the TT should be conceived as a comment on human judgment and could be used as an 'anthropological tool', as a means for research of the factors that affect human judgment. In this sense our proposal is fully aligned with the concept of the TT, aiming actually not at the evaluation of GR3 but at the research on the factors that rule this evaluation.

Thus in this 'singing voice TT' that we propose we will shift our interest from the evaluation of the singing voice 'qualities'("vocalness", "naturalness", "intelligibility" and "expressivity") to the factors that govern the attribution of these qualities. Such factors are: *gender*, *age*, *education*, *culture* and *familiarity*.

Gender is an acknowledged factor for the perception and appreciation of music in general [37], [76], [73], [35], [58], [110]. But also it is a factor acknowledged in the case of perception, appreciation and production of the singing voice [53], [54], [4], [21].

Age is another well considered factor for the perception and appreciation of music [92], [89], [102], [99], [116]. Howard has traced changes in pitching skills and the development of the singing voice during childhood [52], [53].

Education and especially music education has been also widely considered as another important factor

for musical preferences and perception [104], [42], [84], [96]. Similarly, training in singing influences the way one perceives the singing voice [105], [3], [80].

Familiarity and specifically familiarity with technology is expected to be one more factor affecting the interrogators in our 'singing voice TT'. *Stimulus familiarity* in general has been interesting the musicologists since the late 60's [17], [67], [83], [7]. But what about familiarity with music technology? A research carried out in the field of AI ethics gives as the initiative to wonder whether the familiarity with technology can be a crucial factor for the perception of singing voice. As already mentioned Swisher, Dobromir and Chemero organised an 'embodied TT' in which ethical responsibilities were to be attributed or not to the robot AIBO. The results of that TT showed that interrogators with a significant level of familiarity with robotic systems were more willing to attribute ethical responsibilities to AIBO than interrogators with less familiarity [101]. Could our 'singing voice TT' reveal similar effects of familiarity in the attribution of 'naturalness', 'expressivity' or intelligibility'? This is a question that we also intend to explore.

Finally, *Cultural diversity* is one more factor that we wish to examine with our 'singing voice TT', since it is well acknowledged by the researchers of music social psychology [39], [82], [95], [36], [55], [77]. Cultural diversity is also believed to be an important factor for the perception of the singing voice [70], [8], [9].

3.2. Design and goals

3.2.1. Grouping of participants

These five factors in the perception and appreciation of music and singing voice guide our choices regarding the way in which our 'singing voice TT' will be organised. Specifically, the body of participants will be formulated by choices that will be in accordance with these five factors, so it will be consisted of: 1) primary school students, high school students, university students and middle aged people, in order to examine possible judgmental differences due to the age factor 2) Each of these groups will consist of an equal number of females and males so that any existing judgmental differences due to the factor of gender could be explored 3) Some of the university students will be students of the music department of the university of Athens while an equal number of the university students will belong to other departments of the university. In this way we wish to examine possible judgmental differences due to the factor of musical education 4) Some of the students of the music department will be students of the postgraduate program of Music Technology while an equal number will be just students of the music department with no experience of music technology applications. In this way we aim to explore possible judgmental differences due to the factor of familiarity with music technology 5) The factor of cultural diversity will be examined in terms of ethnicity diversity. Therefore the group of primary school students will consist of two sub-groups, one consisting of students with Greek descent and one with students of African descent (therefore coming from a musical culture quite different from the Greek one). We will also try to subdivide culturally all the other age groups of participants but this might prove difficult for purely practical reasons (though this can be part of an extended version of our TT in the future).

3.2.2. Choice of singing voice samples

The above mentioned groups of participants will listen singing voice diphone samples which will be divided to the thee following categories: A) natural samples (i.e. samples of diphones produced by a natural singing voice B) synthetic samples (i.e. samples from the GR3 Greek synthetic singing voice database produced with the MBROLA synthesizer [62] C) natural samples that have undergone digital transformation in the fundamental frequency (Fo). This third group of samples will be used in the fashion of the 'antithetic factors' usually used in most of nowadays versions of the TT. It is an 'antithetic factor' that we hope that will help us explore if there are any judgmental differences due to familiarity with music technology (will students of music technology score higher in understanding the natural source of these samples than people that have no experience of music technology?). Our choice to 'harm' digitally the Fo is guided by our intention to examine the importance of Fo in the perception of singing voice as shown by several researchers [100], [1], [113], [2], [85].

Finally, all these three groups of samples will be subdivided to two subgroups: a subgroup with samples of female singing voice and a subgroup with samples of a male singing voice. With this subdivision we wish to explore further the role of gender in the perception of the singing voice.

4. CONCLUSIONS-PERSPECTIVES

The TT has been one of the most popular topics in AI and Philosophy of Mind. Though, when used as an intelligence criterion, it seems to suffer from serious problems even when the so believed measure of artistic creativity is employed: 'authorship', 'subjectivity of human judgment' (regarding intelligence or artistic value) and 'intentional stance' are problems which in their appearance make us realize that what can be judged in a TT is judgment itself. In this sense, the use of the TT can be altered from a disputed and rather dubious criterion of intelligence to a successful 'instrument' for the introspection of human judgment. Specifically, in this paper we propose a treatment of the TT as a research tool for the introspection of aesthetical judgment. Since no TT has been carried out for the singing voice we intend to explore the ability of TT in being an introspective 'instrument' in the case of singing voice perception and evaluation.

Our proposal for a 'singing voice TT' aims in two directions that we nevertheless believe them to be

interconnected. The first direction is that of a theoretical and finally 'anthropological' interest. Our research aims at the exploration of factors that are widely believed to affect human perception of the singing voice: gender, age, music education, familiarity and cultural diversity. In this sense, it falls within the scope of cognitive musicology and social psychology of music. The second direction is rather practical since the evaluation of GR3 concerns the field of development and improvement of singing voice synthesizers. Thus it concerns the field of music technology. In our view these two directions should not be conceived differently, since imitating the vocal tract could be equal to deceiving the ear and deceiving the ear needs the knowledge of 'how' the ear perceives.

5. REFERENCES

[1] Akagi, M, Iwaki, M, & Minakawa, T, "Fundamental frequency fluctuation in continuous vowel utterance and its perception", *ICSLP98*, Sydney, Vol.4, pp. 1519-1522, 1998

[2] Akagi, M, & Kitakaze, H, "Perception of synthesized singing voices with fine fluctuations in their fundamental frequency fluctuations," *Proceedings of ICSLP2000*, Beijing, vol. III, pp. 458-461, 2000

[3] Andreas, E, *The Voice of singing*, Carl Fischer, New York, 1975

[4] Angus, J.A.S., Howard, D.M., and Welch, G.F. "Singing pitching accuracy in children aged 7 to 11", *100th Convention of the Audio Engineering Society*, pp. 1-12, 1996

[5] Ariza, C, "The Interrogator as Critic: The Turing Test and the Evaluation of Generative Music Systems", *Computer Music Journal*, vol. 33, no. 2, 2009

[6] Aucouturier, J, & Pachet, F, "Representing Musical Genre: A State of the Art", *Journal of New Music Research*, vol. 32, no. 1, pp. 83-93, 2003

[7] Bachorick, J.P., Bangert, M, Larke, K, Berger, J, Rowe, R, Schlaug, G, "Emotion in motion: Investingating the time-course of emotional judgments of musical stimuli", *Music Perception*, vol. 26, no. 4, pp. 355-364, 2009

[8] Barras, M.C. & Gouiffès, A.M. "The reception of overtone singing by an uninformed listener", *Proceedings of the third Conference on Interdisciplinary Musicology (CIM07)* Tallinn, Estonia, 2007

[9] Barras, M.C. & Gouiffès, A.M. "The Reception of Overtone Singing by Uninformed Listeners", *Journal of Interdisciplinary Music Studies*, vol. 2, no. 162, pp. 59-70, 2008

[10] Beardsley, M.C. *Aesthetics from Classical Greece to the Present-A short History*, Mac-Millan, New York, 1968

[11] Bedworth, J, & Norwood, J, "The Turing Test is Dead", *Proceedings of the 3rd Conference on Creativity and Cognition*, Association for Computing Machinery, pp. 193-194, 1999

[12] Belgum, E, et al. "A Turing Test for 'Musical Intelligence'?", *Computer Music Journal*, vol. 12, no. 4, pp. 7-9, 1988

[13] Bloomfield, B.P. & Vurdubakis, T, "Imitation Games: Turing, Menard, Van Meegeren", *Ethics and Information Technology*, vol. 5, no. 1, pp.27-38, 2003

[14] Boden, M.A. *The Creative Mind: Myths and Mechanisms*, Routledge, New York, 1990

[15] Bringsjord, S, Bello, P, & Ferrucci, D, "Creativity and the Turing Test, and the (Better) Lovelace Test", *Minds and Machines*, vol. 11, pp. 3-27, 2001

[16] Bringsjord, S, & Ferrucci, D, *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*, Lawrence Erlbaum, Mahwah, New Jersey, 2000

[17] Cantor, G.N. "Children's 'like-dislike' ratings of familiarized and non-familiarized visual stimuli", *Journal of Experimental Child Psychology*, vol. 6, pp. 651-657, 1968

[18] Chan, N, "Abstract of sound oriented CAPTCHA," *Proceedings of the Workshop on Human Interactive Proofs*, Palo Alto, p. 35 CA, 2002

[19] Churchland, P.M. *The Engine of Reason, the Seat of the Soul*, MIT Press, Massachusetts, 1996

[20] Coats, A.L., Baird, H.S. & Fateman, R.J., 2001 "Pessimal print: A Reverse Turing Test," *Proceedings of the Sixth International Conference on Document Analysis and Recognition*, Seattle, pp. 1154–1158, 2001

[21] Cohen, A.J. "Advancing interdisciplinary research in singing through a shared digital reprository", *Acoustics 2008*, Paris, pp. 3177-3182, 2008

[22] Cohen, H, "A Self-Defining Game for One Player: On the Nature of Creativity and the Possibility of Creative Computer Programs", *Leonardo*, vol. 35, no. 1, pp. 59-64, 2002

[23] Cohen, P.R. "If Not Turing's Test, Then What?", *AI Magazine*, vol. 26, No. 4, 2006

[24] Cope, D, *Computers and Musical Style*, Oxford University Press, Oxford, 1991

[25] Cope, D, "Computer Modeling of Musical Intelligence in EMI", *Computer Music Journal*, vol. 16, no. 2, pp. 69-83, 1992

[26] Cope, D, *Experiments in Musical Intelligence*, A-R Editions, Madison, Wisconsin, 1996

[27] Cope, D, *The Algorithmic Composer*, A-R Editions, Madison, Wisconsin, 2000

[28] Cope, D, *Virtual Music: Computer Synthesis of Musical Style*, MIT Press, Cambridge, Massachusetts, 2001

[29] Dahlig, E, & Schaffrath, H, "Judgments of Human and Machine Authorship in Real and Artificial Folksongs", *Computing in Musicology*, vol. 11, pp. 212-219, 1997

[30] Damper, R.I. "The Logic of Searle's Chinese Room Argument", *Minds and Machines*, vol. 16, no. 2, pp. 163-183, 2006

[31] Dennett, D, *Brainchildren: Essays on Designing Minds*, MIT Press, Cambridge, Massachusetts, 1998

[32] Dennett, D, *The Intentional Stance*, MIT Press, Cambridge, Massachusetts, 1987

[33] Dreyfus, H.D. *What computers still can't do: a critique of artificial reason*, MIT Press, 1992

[34] Dreyfus, H. D. & Dreyfus, S. E. "Making a Mind Versus Modelling the Brain: Artificial Intelligence Back at a Branch-Point", *Artificial Intelligence*, vol.117, no.1, Cambridge, Mass,1988, reprint in Boden, M.A. (ed.): *The Philosophy of Artificial Intelligence*, Oxford University Press, Oxford, New York,1990

[35] Egermann, H, Nagel, F, Kopiez, R & Altenmüller, E, **"**Online measurement
of emotional musical experiences using internet-based methods –
an exploratory approach". *International conference for music perception and cognition (ICMPC)*, Bologna, 2006

[36] Epstein, J.S. "Misplaced childhood: An introduction to the sociology of youth and their music". *Adolescents and their Music:If its too loud, you're too old*, Epstein, J.S. (ed), pp. xiii-xxxiv, Garland, New York, 1994

[37] Folkestad, G & Lindström, B, *Gender and experience and attitudes towards computers and technology*, Department of Education and Educational Research, Göteborg University, 1995

[38] French, R.M., "Subcognition and the Limits of the Turing Test". In Millican, P & Clark, A (eds.): *Machines and Thought. The Legacy of Alan Turing*, vol. I, Oxford University Press, Oxford & New York, 2002

[39] Gans, H.J. *Popular Culture and high culture: An analysis and evaluation of taste*, Basic Books, New York, 1974

[40] Greenberg, B, "Experiments in Musical Intelligence and Bach". In Cope, D: *Virtual Music: Computer Synthesis of Musical Style*, MIT Press, Cambridge, Massachusetts, 2001

[41] Halpern, M, "The Trouble with the Turing Test", *The New Atlantis*, vol. 11, pp.42-63, 2006

[42] Hargreaves, D.J., Messerschmidt, P. & Rubert, C, "Musical preference and evaluation", *Psychology of Music*, vol. 8, pp. 13-18, 1980

[43] Harnad, S, "Minds, Machines and Turing", *Journal of Logic, Language and Information*, vol. 9, no. 4, 2000

[44] Harnad, S, "Other Bodies, Other Minds: A Machine Incarnation of an Old Philosophical Problem", *Minds and Machines*, 1991

[45] Hauser, L, "Look Who's Moving the Goal Posts Now", *Minds and Machines*, vol. 11, pp. 41-51, 2001

[46] Hayes, P & Ford, K, "Turing Test Considered Harmful", *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, vol. 1, 1995

[47] Hiraga, R, et al., "Rencon:Towards a New Evaluation Method for Performance Rendering Systems", *Proceedings of the International Computer music Conference, International Computer Music Association*, San Francisco, California, pp. 357-360, 2002

[48] Hiraga, R, et al., "Rencon2004: Turing Test for Musical Expression", *Proceedings of the 2004 Conference on New Interface for Musical Expression, Association for Computing Machinery*, New York, pp. 120-123, 2004

[49] Hofstadter, D.R. *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*, Basic Books, New York, 1996

[50] Hofstadter, D.R. *Gödel, Escher, Bach: An Eternal Golden Braid*, Vintage, New York, 1979

[51] Hofstadter, D.R. & Dennett, D (eds.), *The Mind's I*, Basic Books, 1981

[52]Howard, D.M. "Qualifying Developmental singing voice changes in children", *1st International Conference on the Psychology and Acoustics of Singing*

[53] Howard, D.M., Angus, J.A.S., and Welch, G.F. "Singing pitching accuracy from years 3 to 6 in a primary school", *Proceedings of the Institute of Acoustics*, vol. 16, no. 5, pp. 223-230, 1994

[54]Howard, D.M. "Variation of Electrolaryngographically derived closed quotient for trained and untrained adult female singers", *Journal of Voice*, vol. 9, no. 2, pp. 163-172, 1995

[55] Huron, D, " Issues and prospects in studying cognitive cultural diversity", *Proceedings of the 8th International Conference on Music Perception and Cognition*, Evanston, pp. 93-96, 2004

[56] Jefferson, G, "The Mind of Mechanical Man", *British Medical Journal*, vol. 1, pp. 1105-110, 1949

[57] Kochanski, G, Lopresti, D & Shih, C, "A reverse Turing Test using speech", *Proceedings of the seventh International Conference on Spoken Language Processing*, Denver, pp. 1357-1360, 2002

[58] Koelsch, S, et.al. "Electric brain responses reveal gender differences in music processing", *Journal of Cognitive Neuroscience*, vol. 15, no. 5, pp. 683-693, 2003

[59] Kugel, P, "Computers Can't Be Intelligent (…and Turing Said So)", *Minds and Machines*, vol. 12, no. 4, pp. 563-579, 2002

[60] Kurtzweil, R, *The Age of Intelligent Machines*, MIT Press, Cambridge, Mass., 1990

[61] Kurtzweil, R, *The Age of Spiritual Machines*, Penguin Books, New York, 1999

[62] Kyritsi, V, Georgaki, A & Kouroupetroglou, G, "A score-to-singning voice synthesis system for the Greek language", *Proceedings of the International Computer Music Conference (ICM'07)*, Copenhagen, Denmark, pp. 216-223, 2007

[63] Lopresti, D, Smith, C & Kochanski, G, "Human interactive proofs for spoken language interfaces," *Proceedings of the Workshop on Human Interactive Proofs*, Palo Alto, CA, pp. 30–34, 2002

[64] Lovelace, A (1842) "Translator's Notes to an Article on Babbage's Analytical Engine". In Taylor, R. (ed.): *Scientific Memoirs: Selected from the Transactions of Foreign Academies of Science and Learned Societies, and from Foreign Journals*, vol. III, pp. 691-731, Richard and John Taylor, London

[65]Loy, D.G. "Connectionism and Musiconomy", *Proceedings of the International Computer Music Conference*, International Computer Music Association, San Francisco, California, pp. 364-374, 1991

[66] Marsden, A, "Music Intelligence and Artificiality", In *Readings in Music and Artificial Intelligence*, Miranda, E.R. (ed.), Routledge, New York & London, 2000

[67] McMullen, P.T. "Music as perceived stimulus object and affective responses as an alternative theoretical framework". In *Handbook of Music Psychology*, Hodges, D.A. (ed), National Association for Music Therapy, Lawrence, Kansas, 1980

[68] Michie, D, "Turing's Test and Conscious Thought". In *Machines and Thought. The Legacy of Alan Turing*, vol. I, Millican, P & Clark, A (eds.), Oxford University Press, Oxford & New York, 2002

[69] Mostow, J & Rich, C, "The Fifteenth National Conference on Artificial Intelligence", 1998 (online at the address: www.aaai.org/Conferences/AAAI/aaai98.php)

[70] Nakayama, I, "Comparative studies on vocal expression in Japanese traditional and western classical-style singing, using a common verse," *Proceedings of ICA 2004*, pp.1295-1296, 2004.

[71] Naor, M, "Verification of a Human in the Loop or Identification via the Turing Test", unpublished manuscript (uploaded on the Internet), 1996

[72] Newell, A & Simon, H.A. "Computer Science as Empirical Enquiry: Symbols and Search". In the *Tenth Turing Lecture, Communications of the Association for Computing Machinery*, 19, Mar.1976, reprint in Boden, M.A. (ed.): *The Philosophy of Artificial Intelligence*, Oxford University Press, Oxford, New York, 1990

[73] O'Neill. S, "Gender and music". In *The Social Psychology of Music*, Hargreaves, D.J. & North, A.C. (eds), Oxford University Press, Oxford, UK, pp. 46-63, 1997

[74] Pachet, F, "The Continuator: Musical Interaction with Style", *Proceedings of the International Computer Music Conference*, International Computer Music Association, California, San Francisco, pp.211-28, 2002

[75] Pachet, F, & Cazaly, D, "A Taxonomy of Musical Genres", *Actes du cogrès RIAO (Recherche d'Information Assistée par Ordinateur) 2000: Content-Based Multimedia Information Access*, Centre des Hautes Etudes Internationales d'Informatique Documentaire, pp. 1238-1246, Paris, 2000

[76] Panksepp, J. "The emotional sources of 'chills' inducted by music", *Music Perception*, vol.13, no. 2, pp.171–207, 1995

[77] Patel, A.D, Inversen, J.R. & Ohgushi, K, "Cultural differences in rhythm perception: What is the influence of native language?", *Proceedings of the 8th International Conference on Music Perception and Cognition*, Evanston, pp. 88-89, 2004

[78] Pearce, M, "The Construction and the Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition", doctoral dissertation, Department of Computing, City University, London, 2005

[79] Pearce M, Meredith, D & Wiggins, G, "Motivations and Methodologies for Automation of the Compositional Process", *Musicae Scientiae*, vol. 6, no. 2, pp. 119-147, 2002

[80] Reid, C, *Voice: Psyche and soma*, Joseph Patelson Music House, New York, 1975

[81]Roads, C, "An Overview of Music Representations". In Roads, C: *Musical Grammars and Computer Analysis*, Olschki, L.S., Firenze, 1984

[82] Roe, K, "Youth and music in Sweden: Results from a longitudinal study of teenagers' media use", Mediapanel, report no. 32, Lunds Universitet, Sociologiska Institutionen, Lund, 1984

[83] Russell, P.A. "Effects of repetition on the familiarity and likeability of popular music recordings", *Psychology of Music*, vol. 15, pp. 187-197, 1987

[84] Russell, P.A. "Relationships between judgments of the complexity, pleasingness an interestingness of music", *Current Psychological Research*, vol. 2, pp. 195-202, 1982

[85] Saitou, T, Tsuji, N, Unoki, M, & Akagi, M, "Analysis of acoustic features affecting "singingnes" and its application to singing voice synthesis from speaking voice," *Proceedings of ICSLP2004*, Vol. III, pp. 1929-1932, 2004

[86] Savova, V, & Peshkin, L, "Is the Turing Test Good Enough? The Fallacy of Resource-unbounded Intelligence", *International Joint Conference on Artificial Intelligence (IJCAI)*, 2007

[87] Saygin, A.P., Cicekli, I & Akman, V, "Turing Test: 50 Years Later", *Minds and Machines*, vol. 10, no. 4, 2000

[88] Searle, J, "Minds, Brains, and Programs", *Behavioral and Brain Sciences*, vol. 3, no. 3, pp. 417-457, 1980

[89] Serafine, M.L., Crowder, R.G. & Repp, B.H. "Intergration of melody and text in memory for songs", *Cognition*, vol. 16, pp. 285-303, 1984

[90] Shieber, S. M. "Lessons from a Restricted Turing Test", *Communications of the ACM*, vol. 37, no. 6, 1993

[91] Shirali-Shahreza, S, Ganjiali, Y & Balakrishnan, R, "Verifying Human Users in Speech-Based Interactions", *Interspeech2001*, pp. 1585-1588, 2011

[92] Shuter-Dyson, R & Gabriel, C, *The psychology of musical ability*, 2nd edition, Methuen, London, 1981

[93] Silva, P, *David Cope and Experiments in Musical Intelligence*, Spectrum Press, Los Angeles, California, 2003

[94] Simon, H.A., *Sciences of the Artificial*, MIT Press, Cambridge, Mass., 1969

[95] Sloboda, J.A. "Empirical studies of emotional response to music", In Jones, M.R. & Holleran, S (eds.) *Cognitive Bases of Musical Communication*, American Psychological Association, Washington D.C., pp. 33-46, 1992

[96] Sloboda, J.A. *The musical mind: The cognitive psychology of music*, Oxford University Press, Oxford, UK, 1985

[97] Soldier, D, "Eine Kleine Naughtmusik: How Nefarious Nonartists Cleverly Imitate Music", *Leonardo Music Journal*, vol. 12, pp. 53-58, 2002

[98] Solis, J, et al., "The Waseda Flutist Robot WF-4RII in Comparison with a Professional Flutist", *Computer Music Journal*, vol. 30, no. 4, pp. 12-27, 2006

[99] Stipp, H, "Musical demographics. The strong impact of age on music preferences affects all kinds of businesses", *American Demographics*, August, pp. 48-49, 1990

[100] Sundberg, J, *The Science of the Singing-Voices* Northern Illinois University Press, 1987

[101] Swisher, N, Dotov, D & Chemero, A, "Ascribing Moral Value and the Embodied Turing Test", 2009, article published online at the web-page: *PhilPapers: Online Research in Philosophy*, at the address: http://philpapers.org/rec/CHEAMV. Also published at the web-page: *CiteSeerX-Scientific Literature Digital Library and Search Engine*, at the address: http://citeseerx.ist.pru.edu/viewdoc/symmary?=10.1.1.142.4055

[102] Tolhurst, G.C., Hollien, H. & Leeper, L. "Listening preferences for music as a function of age", *Folia Phoniatricia*, vol. 36, pp. 93-100, 1984

[103] Turing, A, "Computing Machinery and Intelligence", *Mind*, Vol. LIX, No. 2236, 1950, reprint in Boden, M.A. (ed.): *The Philosophy of Artificial Intelligence*, Oxford University Press, 1990

[104] Valentine, C.W. *The experimental psychology of beauty*, Methuen, London, 1962

[105] Vennard, W, *Singing: The mechanism and the technic*, Carl Fischer, New York, 1967

[106] von Ahn, L, Blum, M, Langford, J & Manber, U, 2001 "The CAPTCHA project: Telling humans and computers apart (automatically)," http://www.captcha.net/., October 2001

[107] von Ahn, L, et al., "CAPTCHA: Using Hard AI Problems for Security", *Advances in Cryptology-Eurocrypt 2003*, International Association for Cryptologic Research, Santa Barbara, California, pp. 294-311, 2003

[108] Wakefield, J.C., "The Chinese Room Argument Reconsidered: Essentialism, Indeterminacy, and Strong AI", *Minds and Machines*, vol. 13, pp. 285-319, 2003

[109] Weinberg, G, & Driscoll, S, "Toward Robotic Musicianship", *Computer Music Journal*, vol. 30, no. 4, pp. 28-45, 2006

[110] Welch, G, et. al. "Musical genre and gender as factors in higher education learning in music", *Research Papers in Education*, vol. 23, no. 2, pp. 203-217, 2008

[111] Whitby, B, "The Turing Test: AI's Biggest Blind Alley?". In *Machines and Thought*, vol. I, Millican, P & Clark, A (eds.), Oxford University Press, 2002

[112] Wiggins, G.A. "A Preliminary Framework for Description, Analysis and Comparison of Creative Systems", *Knowledge-Based Systems*, vol. 19, pp. 449-458, 2006

[113] Yatabe, M, & Kasuya, H, "Dynamic characteristics of fundamental frequency in singing," *Proceedings of the Autumn Meeting of the Acoustical Society of Japan*, 1998

[114] Zdenek, S, "Passing Loebner's Turing Test: A Case of Conflicting Discourse Functions", *Minds and Machines*, vol. 11, 2001

[115] Ziff, P, "The Feelings of Robots", *Analysis*, vol. 19, pp. 64-68, 1959

[116] Zillmann, D & Gan, S, "Musical taste in adolescence". In *The Social Psychology of Music*, Hargreaves, D.J. & North, A.C. (eds), Oxford University Press, Oxford, UK, pp. 161-187, 1997

[117] Zorn, J, "Preface". In Zorn, J (ed.): *Arcana: Musicians on Music*, Granary, New York, 2000