

# The Visceral Voice

## Breath and the Politics of Vocalization

Frank Millward<sup>1(✉)</sup> and Heather Keens<sup>2</sup>

<sup>1</sup> University of Newcastle, Newcastle, Australia  
Frank.Millward@newcastle.edu.au

<sup>2</sup> The Art of Voice, Newcastle, Australia  
heather@theartofvoice.com.au

**Abstract.** There are many factors that influence the sound of a ‘visceral’ or emotive affect vocalization. This discussion develops ideas about how we hear and understand emotive content in the human and the synthetic voice, particularly in relation to affect utterances and nonverbal and paralinguistic vocal gestures. This is important for areas of affective computing and for developing human computer interactive voice processes. Focus is given to how the breath or vocal energy and the affordance of sound, along with socio-political and cultural factors have impact on interactive vocalizations and what that implies for the development of the artificial emotionally intelligent voice of the future.

**Keywords:** Visceral · Voice · Affect utterance · Artificial emotional intelligence · Interactive · Synthetic voice · Paralinguistic · Voice empathy · Voice qualities

## 1 Introduction

The interactive synthetic voice is designed and positioned to respond and function in particular scenarios. The manner and form of a synthetic vocal response may be described in a variety of ways - task activation, confirming, instructional, informative or responsive to particular physical or timed events to name a few. “Human-centered” or “experience-centered” design has ‘opened up deeper investigations of the meaning of affect, emotion, and experience’ [1]. This discussion will look at the role played by political, cultural and social aspects involving interaction where the ‘visceral voice’ or emotive vocal interactive processes are in play. This is particularly relevant in understanding how the empathetic voice can be further developed into the interactive synthetic voice and for exploring the nature of a vocal interactive experience with a free talking artificial emotionally intelligent voice.

Synthetic voices read text and are designed to deliver an emotive position dependent on the environment of delivery. In listening to the instructional voice during a Google Map journey, the synthetic voice is delivered in a neutral voice. The interactive response to the driver taking an alternate route is to remap the alternate route and deliver the subsequent vocal instruction that follows with a similar neutral vocal tone. In a gaming situation however, the use of affect or emotive expressions are freely used to heighten

the emotive response and the sense of a constructed reality, usually presented in a neo-filmic sonic environment where the synthetic voice may be underscored with music to assist in creating an emotive soundscape.

Vocal qualities and attributes can be described and aligned with identifiable emotive states. A voice can be described as happy, sad, angry and so on. These emotional positions can be transferred and represented through the voice in responsive and listening modes. Much of the work to date regarding human technology interaction involving emotion has been multi-modal in approach, using a number of indicators in tandem such as facial, skin response, gestural, pulse, respiration and the like when ‘analysing emotions as an aspect of the user experience’ [2].

Our discussion will focus on a paralinguistic approach involving the sounds that are voiced but generally not included in text-based readings. Affect utterances such as the ‘ums’, ‘ahs’, ‘sighs’ ‘growls’ and ‘breath marks’, indicators of passing emotive states, will be considered in suggesting how these affects might be used to give emotive color to a vocal delivery, real or synthetic. The intention will be to better understand the links between the human voice, the visceral, the breath and the politics of expression in our social, cultural and networked world. To look towards an artificially transacted vocal future when the free speaking synthetic voice may (or perhaps may not) be imbued with the ability to convincingly read and express human emotions using an artificial emotionally intelligent voice.

## 2 The Visceral

The James–Lange theory ‘focuses on bodily change in the muscles and viscera as causing the feeling component of the overall emotional experience rather than being concomitant of it’ [3].

These early theories were typically found to be incorrect or, more commonly, not sufficient in themselves to explain the motivational states in question ... and theories focusing on the central nervous system (CSN) control are now more common. Nonetheless these theories remain about the body’ [3].

There is still some worthy discussion to be had here regarding the ‘visceral voice’. Lavine in his book *In An Unspoken Voice* presents the thought that ‘many of our most important exchanges occur simply through the “unspoken voice” of our bodies’.

The visceral sense is our capacity to directly perceive our gut sensations and those of other organs, including our heart and blood vessels. Most medical texts state that a refined visceral sense is not possible, that “gut feelings” are just a metaphor and that we are only able to feel pain “referred” from the viscera to more superficial body regions. This is ... wrong in fact, without the visceral sense we literally are without the vital feelings that let us know we are alive; it’s our guts that allow us to perceive our deepest needs and longings. [4]

The visceral voice, as an expressive medium, might be seen as coming at the end of a body-located internal communicative process. We feel a physiological sensation. The voice responds with an affect utterance, a ‘sigh’ of discomfort or the ‘scream’ of fear motivated by the intensity of a pain. The lungs support the breath that flows through vocal folds after the brain sends a message to respond. The enactment, a compressed

timeline sequence, starts with a physiological sensation and ends with an affective emotive vocalization. A visceral rhythmic motif is produced as a result of the timing of the sequence of these events. The spectrum of the rhythm being directly dependent on the time period enunciated between the physiological sensation and the end of the affect utterance.

The Attack, Decay, Sustain, Release (ADSR) aspect of this rhythmic/sonic envelope may be considered as a form, as used with music synthesizers when shaping the formation of a musical sound, a fundamental building block in the structuring of a single note, syllabic or phoneme like. The idea of a sonic envelope can be thought of either as a micro and macro descriptor, a single note, phrase or as the envelope of an entire discussion or musical piece. A vocal affect utterance might be thought of as a sonic envelope, a sigh or a cry for example.

A baby's visceral vocalizations, a scream motivated by hunger, may be considered to have a short attack and release time with an extended sustain and decay time, the length of which is contingent upon satisfying the hunger. A baby's visceral voice rhythms might be attributed to a number of motivating physiological causes - pain, discomfort, hunger and so on. The timing of the flow of the envelope develops rhythmic motifs that when repeated provides a way of recognizing or identifying particular visceral vocalizations of the baby. Empathies between mother and child are complex and are commonly acknowledged as formations that commence in the womb as the fetus develops a familiarity and bonds with the sound of the mother's voice [5, 6].

While the suggestion about the rhythmic analysis is speculative it offers an idea for framework development in considering ways in which a synthetic voice might be modeled or programmed, with designed rhythmic cells to reflect coded visceral or affect responses. 'Crying has its internal rhythms, which can be discerned based on an aesthetic sensibility that attends to and interprets sound, tone, intensity, volume, pitch and pace' [7].

The driving notion here lies with the idea that the rhythm of a vocalisation could be seen to be at the centre of political considerations regarding such expression. If we moderate or temper a visceral vocal expression, we may be making a political choice - should I say what needs to be said or what I know will be acceptable to this listener - or perhaps we are confronted with adopting accepted lines of social behaviour, such as if we were to not express a pain through a vocal expression but rather hide the pain and not show that we are hurting.

The nature of the choices for a baby is uninhibited and without censorship, the expression comes straight from the sensation via a direct unimpeded pathway to the voice. For an older person the sensation is more than likely to be tempered with considerations of appropriateness and codes of behavior encouraging a more sanitized vocal expression - particularly when issues involving passionate or viscerally motivated belief are involved - the visceral voice is then tempered and contained, it becomes aware adapts to of social conventions or correctness of behavior. In situations of primal utterance the flow also may be uninhibited such as is fight or flight circumstances.

The visceral utterance may be replaced with an acceptable euphemism. The angry response, "That is impossible!" may become, "Umm ... that will be a challenge". In these circumstances metaphor plays a role, as does personal identity and the connectivity

of that identity to the individual sound of the expressive voice. This supports the Jōemets idea that

as soon as voice becomes verbal, it ceases to be voice and becomes language, conveying linguistic meaning by acoustic means ... in order to study voice in its nakedness, it must be viewed from beneath the verbal and musical decorations that have their own modes of creating meaning and that dominate over the purely vocal means of signification. [8]

It is here that the tensions are to be interrogated, where the clarity of an emotion becomes tempered with linguistic considerations and the expressive emotion pushed into managing the breath, connected to the politic of utterance and verbal expression, vocalization.

The ‘visceral voice’ is something the uninhibited child has little problem with until language is acquired, [8] but is constrained in the adult voice as it speaks to be compliant in a socially and culturally coded language. Life itself could be thought of as a journey that commences in the clarity and resonant qualities of the directness of a visceral, emotive voice to evolve through a number of different stages of sounding, into the matured wisdom of a well-considered vocalization. If a synthetic voice is to be believable and useful in free talking discussions with people, should not this evolutionary and developmental ageing process be incorporated into the artificial emotionally intelligent voice of the future?

### 3 Breath and Voice Quality

Emotive indicators such as anxiety, fear, joy, sadness, anger, surprise, disgust can be variously read as indicated in the sound of a voice, emotive experiences subliminally contained within the vocal sound, not dependent on the linguistic content, delivered by a particular set of vocal qualities or indicators. In early research involving the voice and emotive expression, Scherer suggested that ‘the key to the vocal differentiation of discrete emotions seems to be *voice quality*’ [9].

Since then there has been massive amount of research done in analyzing and understanding the emotive voice [23–25], especially involving fundamental frequency F0, range and pitch and how these indicators are rendered or read as emotive signaling. More recent studies involving voice qualities suggest ‘that there is no one-to-one mapping between voice quality and affect: individual qualities appear rather to be associated with a constellation of affective states, sometimes related, sometimes less obviously related’. [10]

‘Whoa-oa-oa!’ - the opening scream from the song ‘I Feel Good’ by James Brown is an example of a heartfelt voice expressing joy. ‘Whoa-oa-oa!’ is also an example of vocal constriction (pharyngeal constriction). This occurs when the Vestibula folds (false vocal folds) act as a dampener on the true folds, due to excess or under-excess breath flow combined with laryngeal and physical body tension. When the vocal folds (or vocal chords) - the fleshy structures in the larynx that produce the sound of the voice - have been compromised. Instead of vibrating freely, the folds slam together in a compressed space producing a distorted sounding voice. If the voice is consistently constrained in this manner - through continual shouting for example - damage can occur where a callous

or nodules form. As a result, the voice produces a hoarse or ‘raspy’ sound, or sometimes no sound at all.

‘Hoarseness’ is used to describe the sound of a voice with damaged vocal folds, oedema (swelling) caused by constriction or vocal fatigue, or folds in a state of pharyngeal constriction, sometimes the sound is referred to as ‘harsh’ voice. This is one of a number of terms used to describe voice qualities, in speech and singing, including - breathy voice, whispery voice, creaky voice and lax-creaky voice, cry, twang, and for the sung voice; speech quality, opera quality, belt, siren, ‘vocal fry’ and so on. The exact meaning of such terms is often problematic as word descriptions of a vocalized sound can mean different things to different listeners. Nevertheless these terms, as currently developed by voice practitioners and specialists, form a common language for voice quality descriptors and moreover often assume associated emotive qualities.

The idea of character is perhaps more relevant than a distinctive definable emotive quality, in as much as the sound afforded by a ‘hoarse’ or damaged voice brings with it certain notions of a life of experience. A truth, worldliness, a wisdom not present in an emotionally neutral voice. Just as we might consider that a higher pitched (young sounding) voice may project the idea of innocence. Focusing on the sonic quality of a voice, the idea is that the breath drives a number of factors when the voice communicates emotion, mood or attitude. Breath is a fundamental visceral component in producing the sound of a voice. Its function also signals the rhythm of a vocal delivery, as ‘gaps in speech necessary for breathing are governed by syntactic structure of the language’ [11].

The idea of experiencing the rhythmic affordance of sound is relevant in the sense that many emotive utterances and paralinguistic vocalizations come as embedded human knowledge. Further, such sounds are learned as a result of social interaction and the development of codes of behaviour throughout a person’s lifetime.

sounds are not lived in isolation, but experienced through the lived context of social representations that govern how we listen and hear. As Hayes-Conroy and Hays-Conroy (2008:467) [12] remind us: “In the visceral realm, representations affect materially”. Following this lead, the historical weight and orientation of social norms aligned to sounds become part of new intensities, memories and emotions-sounds mobilise visceral mechanisms that help particular political subjectivities to temporally fluoresce. [13]

In our technologically socialized world the mediated and processed real and synthetic voice has shaped the way we hear and perceive the sound of the disconnected and the visually aligned ‘voice’. This is particularly relevant when considering spatial context, consonance, dissonance, rhythmic flow, factors associated with social behavior and what could be described as embedded knowledge – relevant here is embedded sonic knowledge that makes it intuitive for humans to decide on a course of action such as fight or flight as described above, the ability to differentiate between ‘between distressed and other kinds of cry’ [7].

The Scherer model of *time-frequency-energy measures* [9] could be added to and repositioned as - *duration, pitch variation, breath, visceral rhythmic indicators along with a range of considerations to do with social context and cultural understandings* – translated, codified or modelled this becomes a complex array of layers that interact within and between layers in a voice as well as between voices in discussion, nested rhythms - rhythms within rhythms, within rhythms. Fractal formations, when looked

into reveal another within, then another, and another and so on. The connection to be made here is with neuroscience investigations involving the rhythms of the brain where the idea of nested rhythms is associated with notions of consciousness, processes within processes within processes considered as rhythmic units, the rhythms of living, the rhythm of ‘being’.

The *AjoChhand Artificial Brain Building* project is an example of where this idea is being used - ‘we believe that we need a completely new geometric language of nested rhythm to process information in the fractal architecture’ [14]. Other approaches being used such as ‘Neuromorphic Engineering’, seek ‘to build artificial nervous systems which mimic the functions of biological nervous systems’ and aim to ‘replicate the performance of the brain’ and also to contribute to ‘better brain-machine interfaces’ [26]. Feinberg’s article explores

fundamental questions about the nature of consciousness and the nervous system, and attempts to reconcile certain philosophical positions with neuroscience. In particular I focus on the relationship between consciousness and hierarchical neural structure, emergence, scientific reduction and ontological subjectivity [27].

The connections being made between the idea of ‘nested rhythms’ and ‘consciousness’ is where current research is being used to connect the neural with understandings about emotion and its manifestations physical, psychological and verbal [28]. In this neuroscience environment modeling the visceral voice is within its infancy, but nevertheless there are signs that significant connections are being made in understanding the relationship between, consciousness, the visceral, the computational, the developmental and the evolutionary phases of our life cycles and how that might translate to the sound and nature of the artificial emotionally intelligent voice.

## 4 Empathy and Affect Utterance

While Mirror Neuron research is also in early stages of development, there was an interesting observation made about the way in which a monkey responded to a ‘noisy action’ (any action that produces a sound) and later responded to the sound of that action. ‘The results showed that a large number of mirror neurons, responsive to the observation of noisy actions, also responded to the presentation of the sound proper of that action, alone’ [15]. Is there some thread here for speculating that there is a basis for the idea that empathy can be signaled through vocal sound alone, without the confirming view of a facial expression?

In studies where conversations with virtual human voices have been conducted, synthetic emotions have been used to evaluate responses from the perspective of the human interactive experience. In a study conducted by Qu, Brinkman et al., outcomes relevant to this discussion were noted:

The analyses on the data for valence and discussion satisfaction suggest that positive compared to negative synthetic emotions expressed by a talking virtual human can elicit a more positive emotional state in a person, and create more satisfaction in the conversation [16].

This suggests that the virtual speaking voice is more attractive to the human listener when the virtual voice speaks in a positive, supportive or conformational way.

Responding to this research outcome, designers seeking to ‘elicit emotions’ in a virtual voice should concentrate on making the virtual voice able to speak (and listen) to the human in a positive way.

Another outcome noted that ‘participants seemed less satisfied with the conversation when the virtual voice showed negative instead of positive emotions during the listening phase’. What one might expect, most people would prefer to be acknowledged by positive conformational emotive interaction rather than experience a negative presence. This supports the idea that the politic of expression is best framed with positive emotions that present empathy through engagement, the negativity of aggression, fear or anxiety with a virtual or real person of course is much less appealing.

Currently HMM (Hidden Markov Model) based speech synthesis is used to produce stronger emotional readings in synthetic speech where existent speaker emotions are modeled and transposed onto a synthetic voice [17]. The study conducted by Lorenzo-Trueba et al. tested four learned emotions - anger, happiness, sadness and surprise in their ‘emotion transplantation’ project to produce an ‘increase in the perceived emotional strength’ [18] of a synthetic voice – when reading texts. The emotion transplantation method is described as a

method capable of learning the paralinguistic information of emotional speech, control its emotional strength and transplant it to different speakers for whom we do not have any expressive information. We decided to focus on emotional speech as a particularization of expressive speech ... we can expect the transplantation method to be able to support different expressive domains [18].

Changes in the response of the listening and the speaking virtual voice may be enhanced by programming the virtual voice with conformational affect or visceral utterances during the listening and speaking phase; sounds such as the positive ‘Umm’, the confirming ‘ah’, the ‘sigh’ of acknowledgement and so on. The inclusion of such affect signaling in the virtual speaking voice would provide paralinguistic sonic material to make the prosody familiar to the human listener.

Modes of speaking have been drawn through positions on a cultural hierarchy, tribal division or identified as a politic expressed through ways of speaking. For example the affect utterances of particular dialects or accents or cadential forms such as the upward inflection in ending a sentences or ways of speaking that can be identified as a cultural identifier, are now readable in a globally connected socially networked world.

The breath in combination with voice quality identifies emotive sonic attributes in a voice. To know ourselves through how we speak, to identify who we are, what we know and connect to the visceral sense that ‘arises through receptors in the gut’, as an intrinsic part of that knowing. The ‘gut’ feeling vocalized or heard in another person’s voice. ‘The most intimate sense we have of ourselves is through proprioception, kinesthesia and visceral sensation’ [4]. The visceral sensation translated as reaction, the scream, the sigh, the expletive all sustained by the life sustaining breath.

But is it possible for a synthetic or virtual voice to be visceral? Where do we stand with tampering with the line between reality and fiction or comfortable truth, e.g. being supportive or annoyingly positive, reflecting and speaking only what wants to be heard. This begs the consideration of a more profound position put by Boddam-Whetham



‘Voice’ can be taken as a phenomenon that takes quite a privileged place; what it utters are not words or phrases as such, but it is rather a voice that speaks in between our words, it contaminates them. In short, the voice of the friend speaks the truth of our Being-with others, it calls from a relation between us which is arguably beyond both the phenomenological subject *qua* philosophical one. The voice opens us up to each other from in-between in a relation of resonance. [19]

Such thoughts point to questions about how we are to achieve social and political understandings when the metaphoric is crossed within and between the boundaries of conflicting realities, synthetic, conceptual and actual. ‘Heidegger’s radical thought in Being and Time was that sociality is a primordial part of existence’ [19]. How can we harness this idea in the networked world where the sincere virtual voice may need to be distinguished from the number of purpose built or *designed voices*. Might it be as simple as making the political decision to ‘un-friend’ the annoying or dishonorable voice?

## 5 Concluding Questions

How can these ideas be made meaningful for the development and improvement of the free speaking interactive synthetic artificial emotionally intelligent voice of the future? If we are to have meaningful, creative and useful dialogues with machines, with synthetic voices, perhaps we need to consider and understand the possible framework in which such dialogues might take place. What should be the role of the synthetic voice? Should that be something that the human is able to choose as a descriptor of interactive persona? - the helpful voice; the cynical voice; the joker; the advisor; the self-designed voice?

Considering the free speaking, emotively sensitive synthetic voice, we might benefit through being able to discuss the creative development of ideas, complex problem solving, philosophical or critical analyses to name but a few possible interactive discussions. A voice that would effectively communicate and read emotions, visceral, tempered or uninhibited, and respond with a number of different, choose-able degrees where ‘visceral’ and ‘logical’ occupy opposite ends of a complex nested set of spectra. A voice that is able to combine advising, discussing and counseling. A voice that would be able to interact with a person who may be stressed or emotionally unstable,

A new USC study suggests that patients are more willing to disclose personal information to virtual humans than actual ones, in large part because computers lack the proclivity to look down on people the way another human might [20]

The idea of a socially and culturally informed voice, one that matures with a life relationship with a human voice, or might be educated to be culturally, socially and politically informed and be aware that these aspects of the communication process are important. Should the synthetic voice be imbued with such qualities and abilities? Should that include a social, cultural and political intelligence, an identity, and a mindfulness of its power to communicate an emotive presence in conversations with real people, one to one or networked? If such a voice were networked how could this contribute to bringing cultures together? How are we to negotiate the political ramifications of designing such a voice? How would that voice be heard? It is clear that we are always on the cusp of another possible great leap forward. That leap will involve the role the mind plays in direct control or instigation of interactions with technology.



What seems clear is that keyboard/screen-mediated human-computer interaction will be a thing of the past. Communication through brainwaves or other somatic sources will be able to create changes, movements, colors, and sounds in our fully-fashioned future. [21]

How will we manage a world ‘in which computational power is also directed at the emotional and psychological dimensions of existence’ [22]? We can only pose questions that may influence the politic of our socially interactive future, where the relationship between the thought and the utterance speak in genuine tones of respect and peaceful co-operation.

## References

1. Lim, Y.-K., Donaldson, J., Jung, H., Kunz, B., Royer, D., Ramalingam, S., Thirumaran, S., Stolterman, E.: Emotional experience and interaction design. In: Peter, C., Beale, R. (eds.) *Affect and Emotion in Human-Computer Interaction*. LNCS, vol. 4868, pp. 116–129. Springer, Heidelberg (2008)
2. Mahlke, S., Minge, M.: Consideration of multiple components of emotions in human-technology interaction. In: Peter, C., Beale, R. (eds.) *Affect and Emotion in Human-Computer Interaction*. LNCS, vol. 4868, pp. 51–62. Springer, Heidelberg (2008)
3. Cameron, O.G.: *Visceral Sensory Neuroscience: Interoception*. Oxford University Press, New York (2002)
4. Lavine, P.A.: *An Unspoken Voice: How the Body Releases Trauma and Restores Goodness*. North Atlantic Books, Berkeley, California (2010)
5. Michelle, N.A., Dorothy, S.M.: The formation of the maternal-fetal relationship: a reflection on the findings of modern medicine. *Natl. Catholic Bioeth. Q.* **15**(3), 443–451 (2015)
6. Welch, G.F.: Speech and song in childhood: a symbiotic development. *Perspect. Voice Voice Dis.* **12**(2), 7–11 (2002). doi:[10.1044/vvd12.2.7](https://doi.org/10.1044/vvd12.2.7)
7. Hopwood, N.: The rhythms of pedagogy: an ethnographic study of parenting education practices. *Stud. Continuing Educ.* **36**(2), 115–131 (2013)
8. Jøemets, V.: Human voice: its meaning and textuality outside the verbal and the musical. *Semiotica* **2014**(198), 305–320 (2014)
9. Scherer, K.R.: Vocal affect expression: a review and a model for future research. *Psychol. Bull.* **99**, 143–165 (1986)
10. Gobl, C., Ni, A.: The role of voice quality in communicating emotion, mood and attitude. *Speech Commun.* **40**(1), 189–212 (2003)
11. Henderson, A., Goldman-Eisler, F., Skarbek, A.: Temporal patterns of cognitive activity and breath control in speech. *Lang. Speech* **8**(4), 236–242 (1965)
12. Hayes-Conroy, A., Hayes-Conroy, J.: Taking back taste: feminism, food, and visceral politics. *Gender Place Culture* **15**(5), 461–473 (2008)
13. Waitt, G., Ryan, E., Farbotko, C.: A visceral politics of sound. *Antipode* **46**(1), 283–300 (2014)
14. AjoChhand Brain Building Project: *Artificial Brain Building: A New Approach to Induce Another Industrial Revolution: A White Paper*, NIMS [www.anirbanlab.co.nr](http://www.anirbanlab.co.nr), [www.nanobraintech.com](http://www.nanobraintech.com)
15. Rizzolatti, G., Craighero, L.: Mirror neuron: a neurological approach to empathy. In: Changeux, J.-P., Damasio, A.R., Singer, W., Christen, Y. (eds.) *Neurobiology of Human Values*, pp. 107–123. Springer, Heidelberg (2005)
16. Qu, C., Brinkman, W.P., Ling, Y., Wiggers, P., Heynderickx, I.: Conversations with a virtual human: synthetic emotions and human responses. *Comput. Hum. Behav.* **34**, 58–68 (2014)

17. Yamagishi, J.: An Introduction to HMM-Based Speech Synthesis. Technical report, Tokyo Institute of Technology (2006)
18. Lorenzo-Trueba, J., Barra-Chicote, R., San-Segundo, R., Ferreiros, J., Yamagishi, J., Montero, J.M.: Emotion transplantation through adaptation in HMM-based speech synthesis. *Comput. Speech Lang.* **34**, 292–307 (2015)
19. Boddam-Whetham, J.: The authentic voice of truth and the resonance of being. *J. Interdisc. Voice Stud.* **1**(1), 41–55 (2016)
20. Lucas, G.M., Gratch, J., King, A., Morency, L.P.: It's only a computer: virtual humans increase willingness to disclose. *Comput. Hum. Behav.* **37**, 94–100 (2014)
21. Flanagan, P.: 2029 fashion futurism. In: *User Experience*, vol. 13, issue 2, 2nd Quarter (2013). <http://uxpamagazine.org/fashion-futurism/>
22. De Botton, A.: Artificial Emotional Intelligence Will Put Old Heads on Shoulders of All Ages: The World in 2016. In: *Wired*, London, pp. 50–53 (2015). <http://www.wired.co.uk/magazine/archive/2015/11/features/wired-world-2016-alain-de-botton-artificial-emotional-intelligence>. (online version subtitled – Six areas that artificial Intelligence will revolutionise)
23. Scherer, K.R.: Vocal communication of emotion: a review of research paradigms. *Speech Commun.* **40**(1), 227–256 (2003)
24. Bänziger, T., Hosoya, G., Scherer, K.R.: Path models of vocal emotion communication. *PLoS ONE* **10**(9), e0136675 (2015)
25. Koelsch, S., Jacobs, A.M., Menninghaus, W., Liebal, K., Klann-Delius, G., von Scheve, C.: The quartet theory of human emotions: an integrative and neurofunctional model. *Phys. Life Rev.* **13**, 1–27 (2015). doi:[10.1016/j.plrev.2015.03.001](https://doi.org/10.1016/j.plrev.2015.03.001). Epub 7 April 2015
26. Orchard, G., Basu, A., Thakor, N.: Building a silicon nervous system: neuromorphic engineering. *Innovations Mag.* **12**(1) (2013)
27. Feinberg, T.E.: Neuroontology, neurobiological naturalism, and consciousness: a challenge to scientific reduction and a solution. *Phys. Life Rev.* **9**(1), 13–34 (2012)
28. Fingelkurts, A.A., Fingelkurts, A.A.: Mind as a nested operational architectonics of the brain: comment on “Neuroontology, neurobiological naturalism, and consciousness: A challenge to scientific reduction and a solution” by Todd E. Feinberg. *Phys. Life Rev.* **9**(1), 49–50 (2012)