

The Ethics of expressive vocal deep-fakes

Nadia GUEROUAOU & JJ AUCOUTURIER

IRCAM / CNRS, Paris & CRP CHU Lille /LiNC INSERM - FEMTO-ST, UBFC/CNRS, Besançon, Alta Voce SAS

Emotions as signals

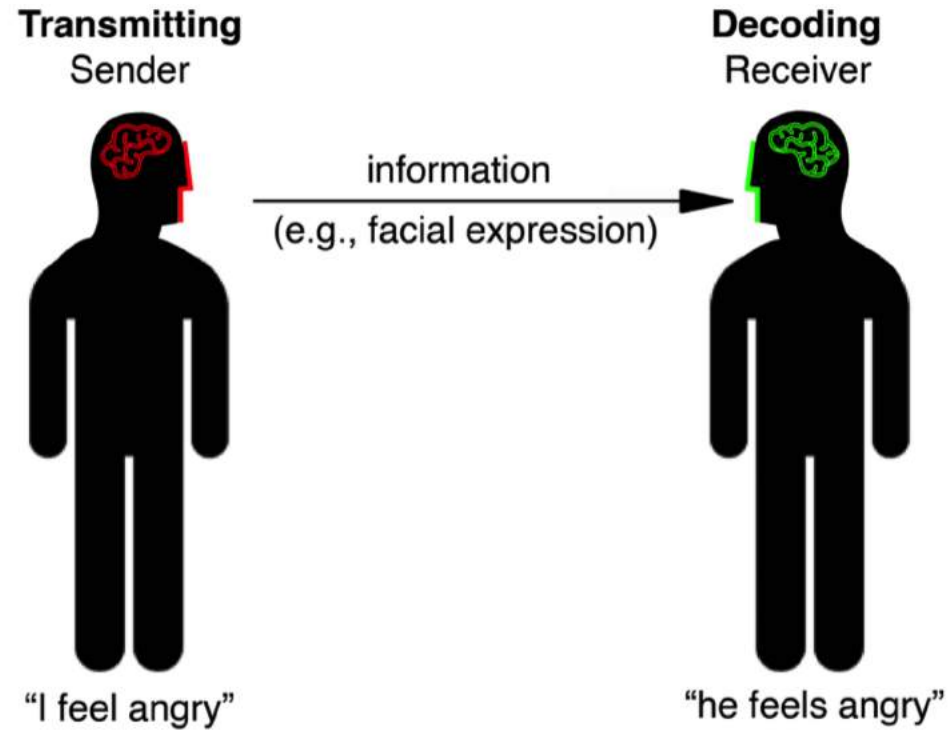
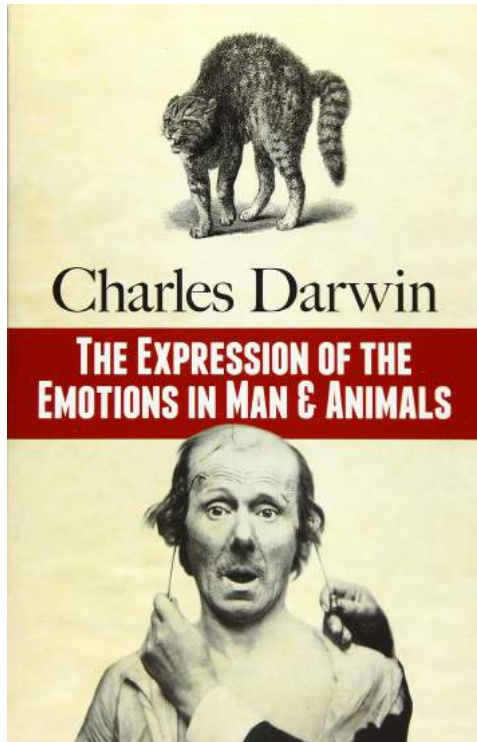


Figure 1: Communication is a dynamic system of information transmission and decoding. Figure from (Jack and Schyns 2015).



//

An easy way to think of a deepfake is like photoshop on steroids. (And powered by artificial intelligence.)

Emergence of Deep-Fakes

Many usages



Emergence of Deep-Fakes

Expressive voice transformations

Enhancing expressivity

Alleviate discriminations

Vocal feedback/Therapy

Cinema

Truthfulness

Augmentation/Fairness

Autonomy/Manipulation

Transparency

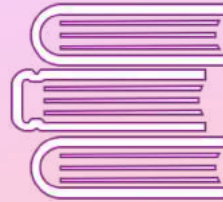
Human identity

Emergence of Deep-Fakes



Controversial media opinions





Moral acceptability of voice transformations?



- Quantify societal attitudes : new technologies & futuristic scenarios
- Comparing different situation variants that may make or break dilemmas
- Assess whether these effects are modulated by individual differences

N=303 online participants

Our cover story
MyVoicePlus™

VoxPressive tech company



24 black mirror-like scenarios



“Imagine you are a depressed patient and you are consulting your physician. You are suffering from the poverty of your social interactions, a symptom of your disease, because your voice is sad and monotonous. Your physician recommends that you use the MyVoicePlus™ vocal prosthesis so that your relatives hear you with a more smiley voice. This will make them more willing to interact with you, and allow you to live better with your condition..”

“Imagine the chairman of a public society consults his life coach because he feels particularly nervous when he has to speak in public. This cause him trouble in his professional activity, which requires him to regularly chair meetings. His coach recommends that he uses the MyVoicePlus™ vocal prosthesis so that his audience hears him with a less tensed voice. This will make him more convincing and appreciated by his colleagues. Imagine that you are a member of this society and that you are regularly in interaction with its chairman while he’s wearing the device..”

“Imagine you are a depressed patient and you are consulting your physician. You are suffering from the poverty of your social interactions, a symptom of your disease, because your voice is sad and monotonous. Your physician recommends that you use the MyVoicePlus™ vocal prosthesis so that your relatives hear you with a more smiley voice. This will make them more willing to interact with you, and allow you to live better with your condition..”

SELF/OTHER

“Imagine the chairman of a public society consults his life coach because he feels particularly nervous when he has to speak in public. This cause him trouble in his professional activity, which requires him to regularly chair meetings. His coach recommends that he uses the MyVoicePlus™ vocal prosthesis so that his audience hears him with a less tensed voice. This will make him more convincing and appreciated by his colleagues. Imagine that you are a member of this society and that you are regularly in interaction with its chairman while he’s wearing the device..”

“Imagine you are a depressed patient and you are consulting your physician. You are suffering from the poverty of your social interactions, a symptom of your disease, because your voice is sad and monotonous. Your physician recommends that you use the MyVoicePlus™ vocal prosthesis so that your relatives hear you with a more smiley voice. This will make them more willing to interact with you, and allow you to live better with your condition..”

THERAPY/ENHANCEMENT

“Imagine the chairman of a public society consults his life coach because he feels particularly nervous when he has to speak in public. This cause him trouble in his professional activity, which requires him to regularly chair meetings. His coach recommends that he uses the MyVoicePlus™ vocal prosthesis so that his audience hears him with a less tensed voice. This will make him more convincing and appreciated by his colleagues. Imagine that you are a member of this society and that you are regularly in interaction with its chairman while he’s wearing the device..”

“Imagine you are a depressed patient and you are consulting your physician. You are suffering from the poverty of your social interactions, a symptom of your disease, because your voice is sad and monotonous. Your physician recommends that you use the MyVoicePlus™ vocal prosthesis so that your relatives hear you with a more smiley voice. This will make them more willing to interact with you, and allow you to live better with your condition..”

ENHANCE POSITIVE/LESSEN NEGATIVE

“Imagine the chairman of a public society consults his life coach because he feels particularly nervous when he has to speak in public. This cause him trouble in his professional activity, which requires him to regularly chair meetings. His coach recommends that he uses the MyVoicePlus™ vocal prosthesis so that his audience hears him with a less tensed voice. This will make him more convincing and appreciated by his colleagues. Imagine that you are a member of this society and that you are regularly in interaction with its chairman while he’s wearing the device..”

Factors variations among vignettes:

- User =participant or other
- Use for therapy or enhancement
- Emotions transformed (positive or negative)
- **Transformation of perception or production**
- **Overt or Covert use**



MEASURES

Judge how morally acceptable it is to use the MyVoicePlus device in such a situation

Judge how acceptable it is to lie to your entourage about using the voice transformation, knowing that this would improve the effectiveness of the device

Judge how acceptable it is that the [doctor/coach] should lie to the user about the voice transformation, knowing that this would improve the effectiveness of the device

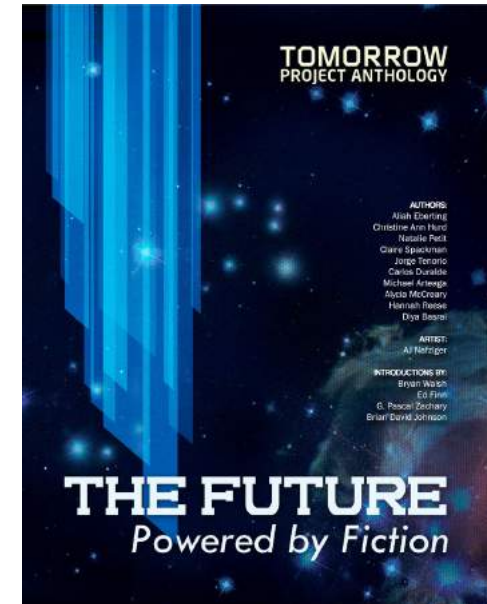
Moral Foundations Questionnaire MFQ

Science Fiction Hobbyism Scale SFH

Vocal deep-fakes are well accepted in the WEIRD population.

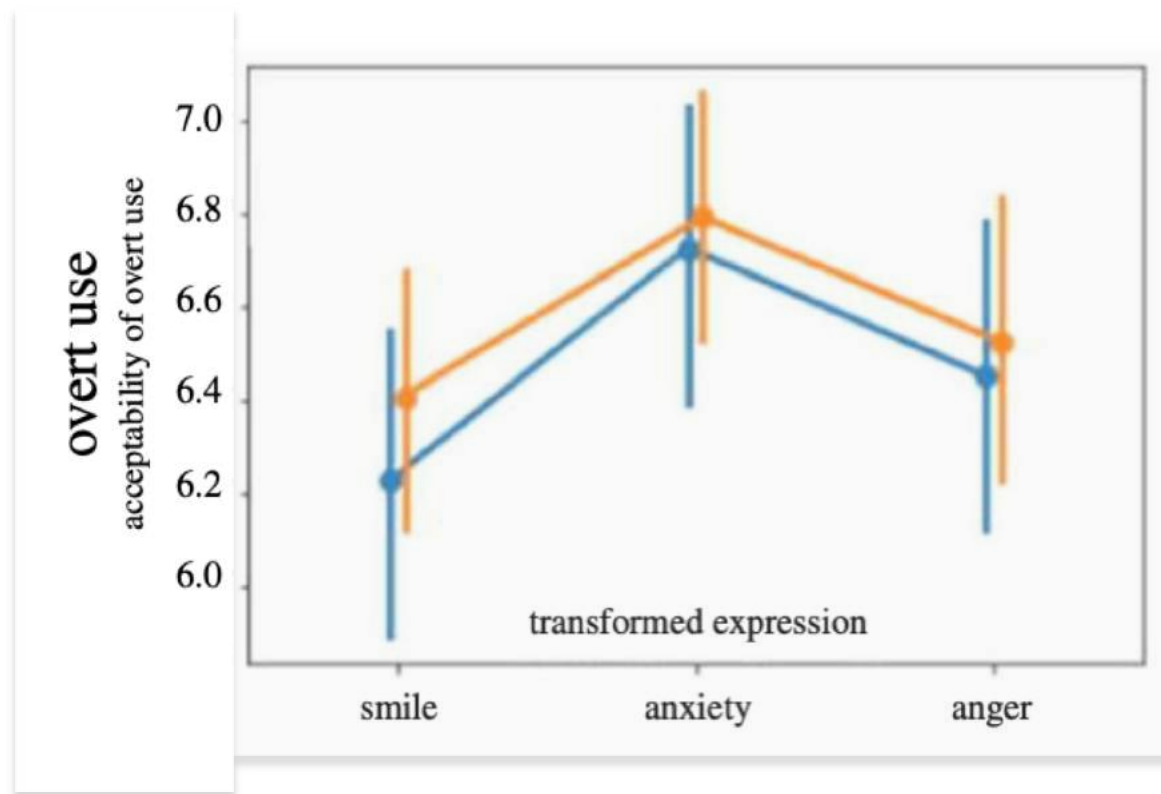
Across situations, the moral acceptability of **overt** vocal transformation was strongly significantly higher than neutral (**M = 6.49 > 5**; one-sample t-test against mid-point, averaging all acceptance scores across vignettes: $t(302) = 146, p < 0.001$).

Familiarity with science fiction



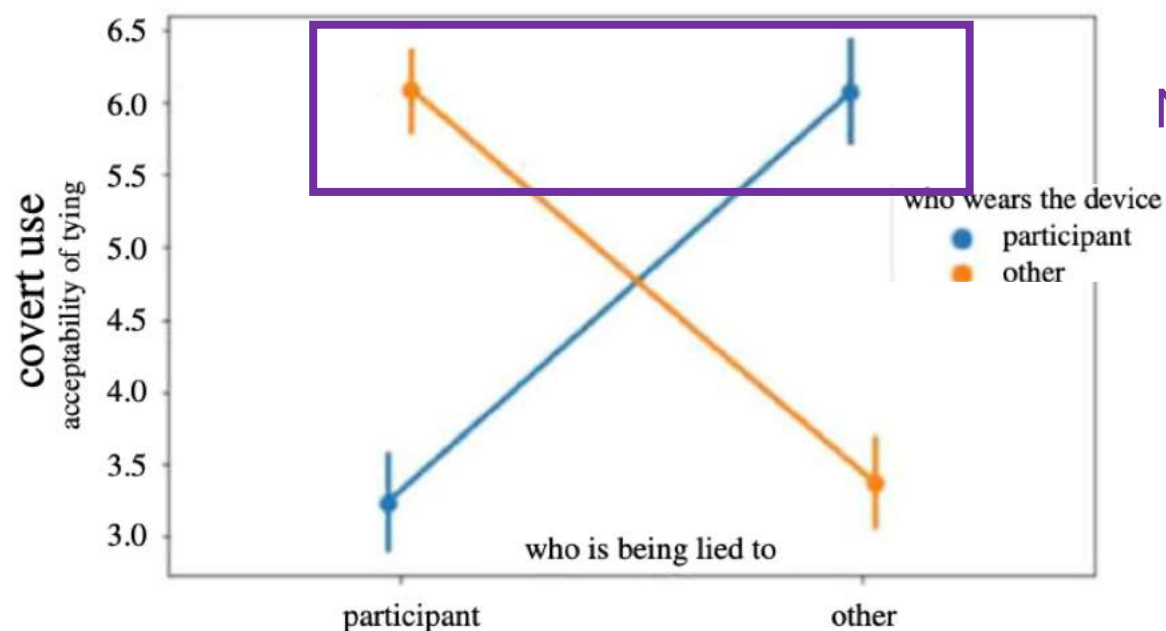
Importance of the **nature of the emotion**:

Curb negative emotions (such as anxiety) > enhance e.g. smiling



Lying about the true purpose of the device:

- the user lies to their interlocutors,
- the device's prescriber's lies to the user.



No ethical issue

NO influence of seeking self-profits.

No evidence of social dilemma SELF \neq OTHER



Conclusions

General acceptability+ Ok Hiding the transformation
-> Massive use of the technology (ZOOM...)



Familiarity with SF
-> Timing of this kind of studies +++

Necessity to inform the population
-> Potential impact on society +++



Change the world with the power of your own voice

PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research



Cite this article: Guerouaou N, Vaiva G, Aucouturier J-J. 2021 The shallow of your smile: the ethics of expressive vocal deep-fakes. *Phil. Trans. R. Soc. B* 20210083. <https://doi.org/10.1098/rstb.2021.0083>

Accepted: 28 July 2021

One contribution of 12 to a theme issue 'Voice modulation: from origin and mechanism to social impact (Part II)'.

Subject Areas:

cognition, behaviour

Keywords:

voice transformation, ethics, deep-fake, moral psychology, emotions

Author for correspondence:

Nadia Guerouaou

The shallow of your smile: the ethics of expressive vocal deep-fakes

Nadia Guerouaou^{1,2}, Guillaume Vaiva² and Jean-Julien Aucouturier^{3,4}

¹Science and Technology of Music and Sound, IRCAM/CNRS/Sorbonne Université, Paris, France

²Lille Neuroscience and Cognition Center (LiNC), Team PSY, INSERM U-1172/CHRU Lille, France

³FEMTO-ST, UBFC/CNRS, Besançon, France

⁴Alta Voce SAS, Houilles, France

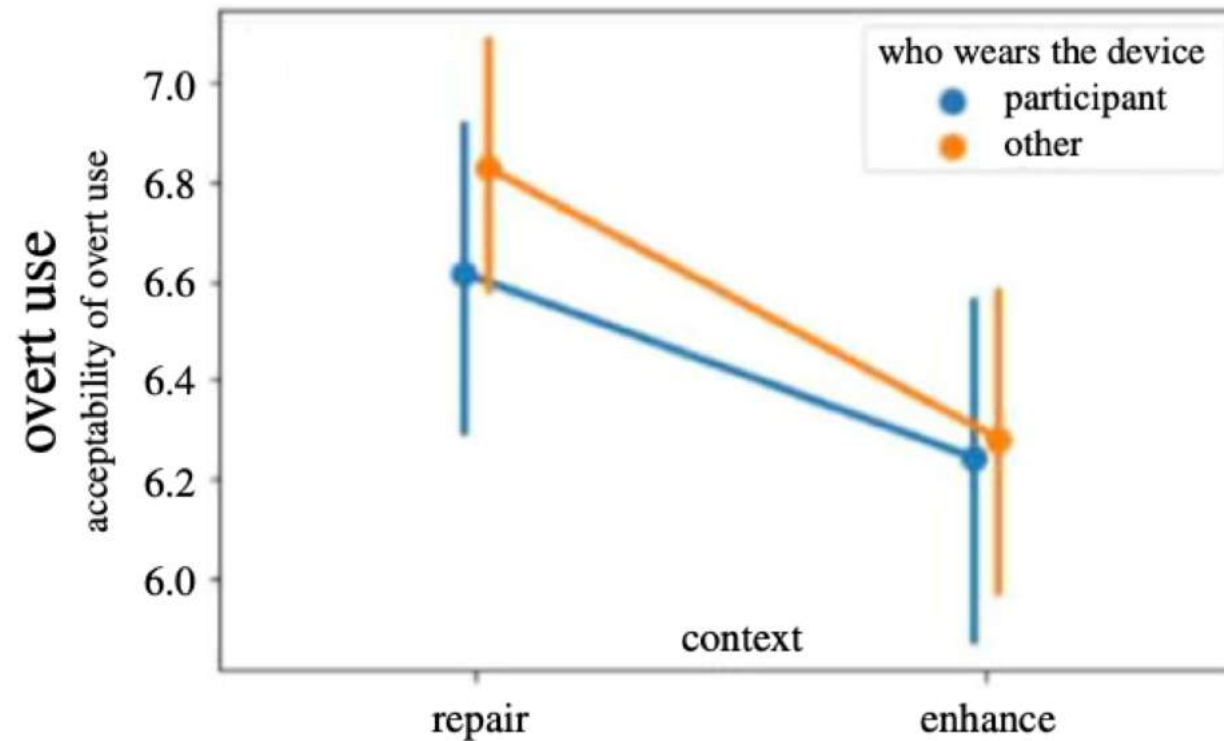
⁴Centre National de Ressource et Résilience (CN2R Lille Paris), Lille, France

NG, 0000-0003-2319-623X

Rapid technological advances in artificial intelligence are creating opportunities for real-time algorithmic modulations of a person's facial and vocal expressions, or 'deep-fakes'. These developments raises unprecedented societal and ethical questions which, despite much recent public awareness, are still poorly understood from the point of view of moral psychology. We report here on an experimental ethics study conducted on a sample of $N = 303$ participants (predominantly young, western and educated), who evaluated the acceptability of vignettes describing potential applications of expressive voice transformation technology. We found that vocal deep-fakes were generally well accepted in the population, notably in a therapeutic context and for emotions judged otherwise difficult to control, and surprisingly, even if the user lies to their interlocutors about using them. Unlike other emerging technologies like autonomous vehicles, there was no evidence of social dilemma in which one would e.g. accept for others what they resent for themselves. The only real obstacle to the massive deployment of vocal deep-fakes appears to be situations where they are applied to a speaker without their knowing, but even the acceptability of such situations was modulated by individual differences in moral values and attitude towards science-fiction.

This article is part of the theme issue 'Voice modulation: from origin and mechanism to social impact (Part II)'.

Therapy > Enhancement



Manipulated **Production** > Manipulated Perception

