Elucidate the Elusive Concept of "Voice" for Interpreters: A Theoretical Model[1]

Caiwen Wang
University of Westminster, UK

Abstract

This paper is an attempt to elucidate the concept of *voice* for interpreters in relation to the equally elusive concept *pleasant voice* for interpreters. The point of departure is that the concept *voice* for interpreting has to do with the physical properties of a speaker's voice, which may lead to the effect that a speaker's voice is heard as pleasant or unpleasant by a listener, depending on how a speaker uses or deploys these physical properties. The paper employs an interdisciplinary approach to reviewing relevant literature and shows that for better interpreter education and interpreting assessment, there is a need to unravel, and unify existing understandings of the concept *voice*. A new definition is therefore proposed. The new definition consists of a cluster of suprasegmental features resulted from supralaryngeal and laryngeal activities and incorporates in what are traditionally known as fluency features in interpreting. The paper goes on to discuss the potential benefits and implications of the newly proposed definition for both interpreter training and interpreting studies.

Keywords: voice; pleasant voice; interpreting; interpreter education; interpreting assessment

Introduction

This paper seeks to define the concept *voice* for interpreter education and interpreting assessment for it to be practical and helpful for assessing whether or not an interpreting voice is pleasant.

We use the term *voice* literally to refer to both the "sound produced by vertebrates by means of lungs, larynx, or syrinx, *especially***:** sound so produced by human beings" and the "expiration of air with the vocal cords drawn close so as to vibrate audibly" (Merriam-Webster Dictionary, 2020; italics original). Our concept of *voice* is therefore different from the

[1] The author is indebted to the two anonymous reviewers and the Editor for their extremely kind and constructive feedback. All errors remain the author's.

also widely studied concept *voice* in Translation Studies, such as in Hermans (1996) and Alvstad *et al* (2017), where *voice* refers to "translators' intervention", as "an index of the Translator's discursive presence" (Hermans ibid., p. 27), or the "individual or collective conceptions and attitudes" expressed by the agents who are "involved in the long and often arduous translation process", namely, those who "read and shape translations – authors, publishers, translators, editors, copy editors, critics, librarians, and 'non-professional' readers". (Alvstad *et al*, ibid., pp. 3-4).

In literature related to interpreters' voice, *voice* has been discussed more in terms of what is known as *right/pleasant voice*, without the two being explicitly distinguished from each other. For example, in her paper "The speech behaviour of interpreters", Horváth (2017, p. 223) states that "[a] right voice leaves a positive voice image, while a wrong voice a negative one", where the difference between what constitutes the "voice image" and what leads to a "positive/negative" voice image could have been spelt out. There have also been survey studies on how important interpreters' pleasant voice is perceived by interpreters and interpreting service users, e.g., Bühler (1986), Kurz (1993), but relevant notion on what *pleasant voice* is not available. In the current paper, we take the stance that the rightness or pleasantness of one's voice is the effect, or impact, or consequence of the way of a speaker/interpreter using their voice properties.[2] Such an effect is recognised or judged from the perspective of a listener. Depending on the way a speaker/interpreter uses their voice when vocalising, their voice may be heard as right/positive/pleasant, or wrong/negative/unpleasant. To achieve a certain effect, such as a pleasant voice, one needs to learn how to vocalise, namely, how to use one's voice or deploy the physical properties of one's voice by means of a range of supralaryngeal and laryngeal activities, so that the impact or consequence from these supralaryngeal and laryngeal activities will lead to the intended effect. It follows that there is a need to delineate a set of features for voice, in order for a specific voice effect to be fully appreciated, or for one to make targeted improvements when one fails to achieve a certain effect due to incorrect or improper ways of deploying one's voice properties. This relationship between voice, the way of using voice and thereafter the voice effect can be represented in the diagram below:

[2] To be concise and gender-neutral, we use the plural pronoun 'they/their' in this paper where appropriate when we refer to a speaker or an interpreter.

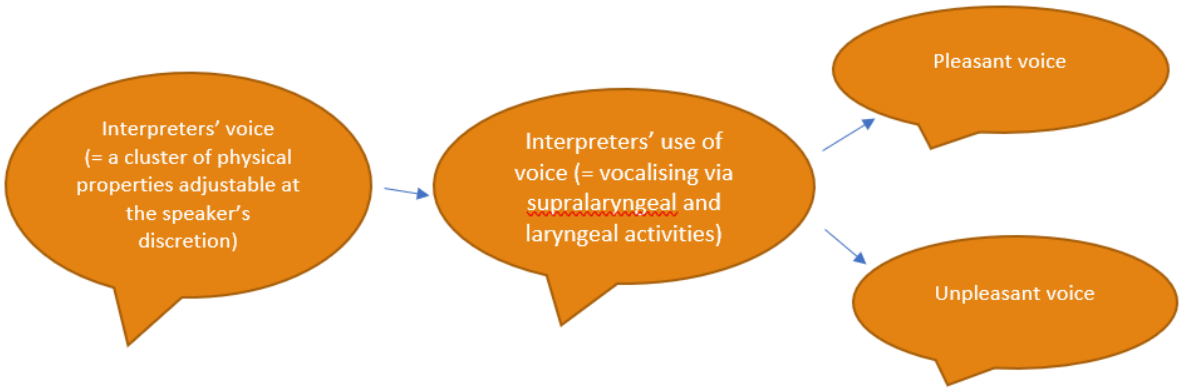

Fig.1 Relationship between voice and (un)pleasant voice

Voice is therefore the fundamental concept for the meaningful discussions of its effects. In the following, we will use Fig. 1 as the reference to show that existing literature has used a bunch of voice-related concepts but have not provided a formal definition for them, which has led to widely different conceptions of their interrelationships, and that where there appears to be some relevant notion, this is not easy to apply to interpreter education or interpreting assessment. We specifically ask the below questions while reviewing the existing literature:

1) Is the relevant literature on voice?
2) Is the relevant literature on using voice?
3) Is the relevant literature on the effect or consequence of one using their voice, i.e., pleasant or unpleasant voice?
4) Is there any definition/notion on what the concept in question is?

In what follows, we examine the existing literature in both interpreting studies and related media studies. The reason for this interdisciplinary approach is that professional bodies such as the International Association of Conference Interpreters (AIIC) and interpreting scholars such as Kurz and Pöchhacker (1995), Pöchhacker (2011) and Nolan (2012) recommend that interpreters look up to newsreaders or commentators for the latter's voice and/or public speaking skills, whilst Schweda-Nicholson (1985) and Cecot (2001) took pains to particularly look into interpreting students' performance in terms of public speaking skills correlated with voice.

Additionally, the AIIC states the following when advising on media interpreting (1999):

> when working for TV or radio, an interpreter's style and delivery need to be particularly smooth and clear, regardless of the original. This is so because TV and radio audiences

are accustomed to the well-trained voices of newsreaders and commentators and do not understand or appreciate the very different demands made of interpreters.

Guidelines like the above reflect two interpreting scenarios: where an interpreter needs to deliver as "clearly and effectively" as a source speaker, and where an interpreter needs to be "smooth and clear, regardless of the original". In the latter case, which implies that a source speaker may not be smooth or clear enough in speaking, an interpreter, who is expected by the AIIC to deliver in a smooth and clear manner beyond the source speaker's alter ego, is compared to newsreaders and commentators, who have a well-trained voice. This analogy certainly makes the concept *voice* easier to understand (even though elusively because "well-trained voices" is a voice effect to be achieved following voice training, which presumes the knowledge of what *voice* is), as in our everyday life we have all heard the voice of a newsreader or commentator.

The AIIC's stance that interpreters' voice needs to be as good as that of newsreaders and commentators "regardless of the original" echoes Collados Aís's view that interpreters should go beyond the "ghost role" (1998, p. 336) regarding intonation. This also highlights the special environment where interpreters work. This is that there is a speaker factor that influences interpreters' voice effect. In order that interpreters can break away from the influence of a speaker who does not speak clearly and smoothly, it becomes even more important to elucidate the concept *voice* to help interpreters consciously monitor the voice features they need to watch out for while interpreting.

Elusiveness of *Voice* in Interpreting Literature

Table 1 presents the interpreting literature to date in relation to the concepts *voice*, *pleasant voice* and/or *unpleasant voice*. Only positive answers to our four questions for literature review are indicated.

Table 1 existing interpreting literature surrounding *voice*

Author(s)	on voice?	on using voice?	on the effect of voice being (un)pleasant?	any definition/notion regarding the concept in question?
Bühler (1986)			Yes	
Kurz (1993)			Yes	
Moser (1996)			Yes	

Collados Aís (1998)			Yes	Yes: Vivid/monotonous intonation was studied separately from instead of as part of pleasant voice.
Chiaro & Nocella (2004)			Yes	
Lambeau (2006)	Yes	Yes	Yes	Yes: voice amounts to acoustic features consisting of volume and pitch; the author suggested two ways for interpreters to have a good interpreting voice, one being looking after one's voice for it to be healthy, and the other being talking to a sound technician or voice coach.
Perng (2006)	Yes			Yes: the author proposed that interpreters' voice be assessed in terms of "control, enunciation, intonation, projection, assurance" (p.198).
Zhang & Ke (2008)			Yes	Yes: Vivid/monotonous intonation was studied separately from instead of as part of pleasant voice.
Liu, Chang & Wu (2008)	Yes			
Pöchhacker & Zwischenberger (2010)			Yes	Yes: Vivid/monotonous intonation was studied separately from instead of as part of pleasant voice.
Iglesias (2013)			Yes	Yes: interpreter users' perception of pleasant voice was found to be "related mostly to prosodic features, and in particular intonation (pitch contour, pitch movement and pitch direction), fluency (speech rate and pause pattern), and diction/articulation. Tone of voice and volume were mentioned much less frequently"; "no reference was made to voice quality or the timbre of the voice and its spectral features" (ibid., p. 107).
Harris (2015)		Yes	Yes	Yes: publishing on the website of the American Translators Association, Harris mentioned two distinctive cases before he discussed six voice registers, one in which his interpreting voice was complimented by a lady who did not know his language and the other in which an interpreting student failed interpreting exam due to his "dreary voice", which in the examiners' view needed a voice coach's help (p. 2); loudness/volume was mentioned as having to do with pleasant or unpleasant voice; projection and endurance were mentioned as having to do with pleasant or unpleasant voice for oratorical interpreting (The six voice registers are the voice for chuchotage, microphone interpreting, telephone

				interpreting, dialogue interpreting, court interpreting and oratorical interpreting).
De Gregoris (2016)	Yes			Yes: an assessment model for TV simultaneous interpreters' voice was proposed for the first time. the sub-categories for voice are: articulation, hesitations, speed of speech, melodious/monotonous voice, same melody repeated, sweet/aggressive voice, active/self-defeating personality of the interpreter, comprehensive voice, expressive voice and credible voice (pp. 298-299). Prosody was separated from voice and was categorised into: audible breaths, silent pauses, natural/non-natural syntax, simple/complex sentences, melodious/monotonous voice, sweet/aggressive voice (p. 298).
Horváth (2017)	Yes	Yes	Yes	Yes: voice, intonation, etc. were subsumed under the umbrella term speech behaviour of interpreters; the importance of correct stress, intonation, speed and voice quality were discussed; the concept *positive voice image* refers to the effect of the whole product package of interpreting, consisting of content, language and non-verbal speech behaviour.
Wang & Huckvale (2018)			Yes	Yes: An individual pleasant female voice was examined in comparison to female voices reading speeches and female voices giving spontaneous speeches. Pleasant voice was associated with the acoustic properties of voice. key findings: (i) the pitch of the pleasant voice was slightly raised when compared to female average for both groups of reference speakers; (ii) the pitch range of the pleasant voice was typical of that of read speeches; (iii) the speaking rate of the pleasant voice was slower than that of the reference speakers, (iv) pauses in the pleasant voice were typical of those found in read speeches, and (v) the pleasant voice displayed controlled vividness of intonation.
Su (2019)	Yes			

As is clear from Table 1, *voice* has not been discussed much in existing interpreting literature when compared with *pleasant voice*, neither *voice* nor *pleasant voice* has been provided with a

formal definition (e.g., Bühler, 1986; Harris, 2015; and Su, 2019), and the understandings of the two concepts are not consistent (e.g., Perng, 2006; De Gregoris, 2016 and Horváth, 2017).

Of special note is the study respectively by Perng (2006), De Gregoris (2016), and Su (2019). These studies are specifically on the assessment of interpreters' voice, thus are closely related to the research aims of the current paper. It is evident that Perng and De Gregoris had very different sub-categories for *voice*: whilst Perng related voice to its physical properties, De Gregoris's subcategories sometimes referred to the physical properties of voice, such as "articulation,*"* but sometimes referred to the various effects of one using one's voice, such as "sweet/aggressive voice,*"* "expressive voice*"* and "credible voice*"*. Also in De Gregoris, *voice* and *prosody* on the one hand were separated from each other but on the other hand overlapped in their actual sense. Furthermore, the concept *prosody* in question included "natural/non-natural syntax, simple/complex sentences*"*, which for interpreting assessment is more often assessed in a *language* rubric and consequently runs the risk of double or multiple penalties. That De Gregoris included "active/self-defeating personality of the interpreter" in the concept *voice* is also debatable. This is because in interpreting it is the quality of an interpreter's voice in relation to interpreting that matters rather than the interpreter's personality, and more importantly because "the interpreter changes voice identity the moment they take on their interpreter hat: they identify with their professional role a different voice and different speech behaviour from their everyday speech characterised by their personality-specific speech habits" (Horváth 2017, p. 232).[3] With regard to Su (2019), no definition of *voice* was provided, and in the assessment criteria provided, *volume* was isolated out of *voice*. Since this study aimed to guide interpreting students' self-study, one would appreciate a greater extent of clarity of the voice criterion to benefit their users.

In view of this, we argue that there is a need to clearly define both *voice* and the voice effect known as *pleasant voice* so that the relationship between the two concepts can be made explicit, thus avoiding variations in perception. Specifically, we argue that voice should be seen as a cluster of physical features which individually and/or collectively contribute to various voice effects among which pleasant and unpleasant voice are relevant to interpreter education and interpreting assessment.

[3] It is, however, possible that interpreters with certain personality traits tend to have a pleasant voice, just like translators having a certain personality are better at literary translation, as found in Hubscher-Davidson (2009). But this would be a different research project.

Moving to *pleasant voice*, this concept was first used by Bühler (1986) in her what is often cited as a seminal study, where a set of sixteen linguistic and extra-linguistic criteria for assessing interpreters' interpreting quality were employed --- with none defined --- to survey professional interpreters' view of their relative importance. In the absence of descriptors designated to *pleasant voice* and other assessment criteria, several studies (e.g., Kurz, 1993; Moser, 1996; Collados Aís, 1998; Chiaro & Nocella, 2004; Zhang & Ke, 2008 and Pöchhacker & Zwischenberger, 2010) followed a similar or identical line of inquiry. These studies on the one hand have shown the importance of a pleasant voice for interpreters, but on the other hand have also brought about new issues needing further investigation. For example, there is the need for a definition of each of Bühler's criteria as without a definition, research subjects' understandings of the criteria, including *pleasant voice*, are likely to be different, an issue that Bühler herself also raised and that was supported by the findings from Iglesias (2013). For another example, there is the need to explicitly categorise *intonation* into *voice*, and accordingly *vivid intonation* into *pleasant voice*. While asking their research subjects to assess pleasant voice, Collados Aís (1998), Zhang and Ke (2008), and Pöchhacker and Zwischenberger (2010) added a category called vivid intonation but did not provide a definition or explanation for the two concepts. We feel this approach gives the impression that vivid intonation is not part of pleasant voice and accordingly intonation is not part of voice, although it is not clear whether or not the researchers in question actually held such a view.

It is worth noting that in Iglesias (2013, p. 107; also see Table 1), interpreting service users' perception of interpreters' pleasant voice made no reference to "voice quality or the timbre of the voice and its spectral features". We believe this would be a useful and appropriate approach for interpreter education and interpreting assessment. According to Laver (2009, pp. 9-10; italics original):

> voice quality derives from two distinctive factors in vocal performance. The first of these is to do with the nature of the individual speaker's own vocal apparatus. The particular anatomy of the speaker constrains his voice quality by the effect of such physical features as the dimensions, mass and geometry of his vocal organs. Thus, organic features such as the length of his vocal tract, the size of his tongue, velum, pharynx and jaw, the shape of his laryngeal structures and the volume of his nasal cavity, will all contribute their effect to the overall quality of the speaker's voice.

> The second factor is to do not with the nature of the vocal apparatus at a speaker's disposal, but the use to which he puts it. Each speaker, as part of habitual style of speaking, tends to use particular settings of his vocal apparatus. … Since these phonetic setting features are all by definition a matter of a mode of control of the muscular apparatus for speech, then all normal speakers should be able to learn to imitate the articulatory basis of the settings, and to recognize their auditory correlates.

The interpreting service users' no reference to "voice quality or the timbre of the voice and its spectral features" in Iglesias (ibid.) very much reflects the first voice factor in the quote above. Indeed, when we assess interpreters' voice quality, we are not assessing singers, who are supposed to have a voice quality in the first sense in Laver's quote above. We should instead assess interpreters' voice quality on the basis of whether or not their voice would make them effective communicators or public speakers, namely, in the second sense in Laver's quote above.

As such, we take Horváth's stance (2017) that interpreters' voice quality has to do with voice production and the process of voice production itself (which is similar to what Laver refers to as the second factor in the quote above for vocal performance), as below:

> Professional interpreters need to be familiar with the speech organs involved in voice production and the process of production itself: not only for the purpose of interpretation but also for long term health reasons. Interpreters, like other public speakers, should have good quality voice as in our ever increasingly competitive profession it might be voice quality that sets them apart from their competitors (ibid., p. 228).

This aside, Horváth ascribes content, language and non-verbal speech behaviour under the umbrella term "positive voice image" (ibid., p. 223; see also the Introduction section). In the context of and for the purpose of interpreter education and interpreting assessment, we feel it would be more helpful to separate the voice criterion from the content and language criterion, respectively, in order to assess interpreters' voice image.

Among the studies presented in Table 1, two terms deserve highlighting. One is intonation. This was studied along with pleasant voice by Collados Aís (1998), Zhang and Ke (2008), and Pöchhacker and Zwischenberger (2010) in terms of its effect of being vivid and being monotonous, respectively, albeit without a definition. By contrast, intonation is explicitly

treated as a voice property by the interpreting service users in Iglesias (2013), though with no definition, either. We align with the interpreting services users in Iglesias and regard intonation as an important voice property.

The other important term is tone. This was touched upon in Nolan (2012) and Iglesias (2013) but was not defined. Following Besson *et al.* (2005), we take *tone* as referring to "a means by which the speaker implies his or her attitude to the message", and "also a means by which he seeks a reaction from the hearer". In other words, tone has to do with emotional regulation, for the purpose of appealing to certain emotional effect. In a political debate, for instance, the tone of voice is likely to sound rousing, whereas on television, news is communicated in a more factual tone.

From Table 1, it can also be seen that interpreting literature on using or deploying voice properties for the pleasant voice effect is fairly sparse. The only works in relation to this are Lambeau (2006), Harris (2015) and Horváth (2017). It is interesting that Lambeau and Harris, who are both seasoned interpreters, suggest using a voice coach to achieve a pleasant voice. Since voice coaches are experts in voice training, it would be ideal to have their help in interpreter education. However, due to the high cost for voice coaching (Trewin, 2003), we do not think this approach is realistic for interpreter training. In our view, the fact that a voice coach will be able to help because they know individuals' voice problem(s) indicates that interpreting tutors will be able to help, too, as long as the assessment criteria for voice are clear.

Elusiveness of *Voice* in Literature on Public Speaking

Table 2 presents the existing mass media literature related to voice and/or (un)pleasant voice and the existing interpreting literature where interpreters' voice or (un)pleasant voice is compared to that of people working in mass media. As with Table 1, only positive answers to our four questions for literature review are indicated.

Table 2 Existing literature on interpreters' *voice* as part of public speaking

Authors	**on voice?**	**on using voice?**	**on the effect of being (un)pleasant?**	**any definition/notion regarding the concept in question?**
Schweda-Nicholson (1985)	Yes	Yes	Yes	Yes: volume of speech and intonation were problems for interpreting students' speech delivery.
Kurz & Pöchhacker (1995)	Yes		Yes	Yes: Media interpreters' voice was compared to a newsreader's or commentator's.

AIIC (1999)	Yes	Yes	Yes	Yes: interpreters were advised to watch out for their volume of voice; interpreters' voice was compared to "the well-trained voices of newsreaders and commentators".
Cecot (2001)	Yes	Yes	Yes	"Voice control, namely loudness, intonation, diction, speech rate and, obviously, pauses are tools at the disposal of public speakers" (p. 68).
Harvey (2003)	Yes		Yes	Yes: according to the author, emotions affected interpreters' voice in interpreting.
Kurz (2003)	Yes		Yes	Yes: according to the author, emotions affected interpreters' voice in interpreting.
Trewin (2003)	Yes	Yes	Yes	Yes: specific suggestions were made for achieving a good presenting voice, as below: • accent: to do with pronunciation. "The limit must be that any accent that interferes with intelligibility is unacceptable" (p. 24) • a calm and relaxing voice • breathing: Control breathing correctly. • projecting: Speak up but not shout. "At the same time you must sound natural" (p. 33) • diction: "Do not be embarrassed to use your mouth to shape words. Make your lips frame the syllables accurately and don't rush. Speak out strongly and clearly" (p. 34) • pace • emphasis and intonation
Valero-Garcés (2005)	Yes		Yes	Yes: according to the author, emotions affected interpreters' voice in interpreting.
Besson *et al.* (2005)	Yes		Yes	Yes: according to the authors, emotions affected interpreters' voice in interpreting.
Russo (2005)	Yes		Yes	Yes: media interpreters' voice was compared to a newsreader's or commentator's.
Lu (2007)	Yes		Yes	Yes: according to the author, intonation was important for good speech delivery.
Ma (2010)	Yes		Yes	Yes: according to the author, volume was important for good speech delivery.
Amato & Mack (2011)	Yes		Yes	Yes: media interpreters' voice was compared to a newsreader's or commentator's.
Bontempo & Malcolm (2011)	Yes		Yes	Yes: according to the author, emotions affected interpreters' voice in interpreting.
Nolan (2012)	Yes	Yes	Yes	Yes: tones and emotions affected interpreters' voice in interpreting; recording was the method for one to correct these.
Gilles (2013)	Yes	Yes	Yes	Yes: volume is a voice feature; good intonation is a public speaking skill; interpreters need to "speak, when interpreting, like a competent public speaker

				giving their own speech" (p. 99); breathing exercises helps for a better voice.
Wang and Mu (2013)	Yes		Yes	Yes: intonation and volume were important for speech delivery.

As with the case of the existing interpreting literature presented in Table 1, in the existing literature related to public speaking, for both interpreting and media studies, a formal definition for *voice* is generally not available. Although more literature talks about voice when discussing voice effects, relevant understandings are elusive. It is also unclear as to whether it is a certain voice property or a cluster of voice properties that constitute voice. Where more than one voice feature is drawn upon for a particular voice effect, it is unclear whether or not they individually and/or collectively contribute to the voice effect in question. For example, when Gilles (2013) advises interpreters on some breathing exercises, he says that this is to "promote better voice quality" (ibid., p. 99) and that "[v]oice coaches will ask you to do this same exercise exhaling" (ibid., p. 100). To our understanding, exhaling is indeed important, but it may not bring about a pleasant voice alone. In a similar vein, Bontempo and Malcolm (2011), and Wang and Mu (2013) factored in one or two voice properties for delivery quality, but it is not clear if the one or two factors are the only factors responsible. In the current paper, we believe that there are additional factors contributing to a voice effect.

Schweda-Nicholson (1985) and Cecot (2001) are two empirical studies on public speaking skills for interpreters. Schweda-Nicholson identified volume of speech and intonation as "common" delivery problems with consecutive interpreting students (1985, p.149). Cecot compared simultaneous interpreters to public speakers, and stated that "[v]oice control, namely loudness, intonation, diction, speech rate and, obviously, pauses are tools at the disposal of public speakers" (ibid., p. 68). Both Schweda-Nicholson and Cecot hold the view that when an interpreting problem is clearly shown to students, it is easier for the latter to take remedial actions. We share this view and are committed to clearly defining *voice* for interpreters, in order for the concept to be a useful didactic tool for assessing interpreters' voice. Unlike Schweda-Nicholson and Cecot, however, we propose that delivery features (such as pauses and speech rate) are part of voice features, rather than the other way around, for the reason that these features are non-verbal and are part of "voice production" (Horváth, ibid., p. 228) resulted from supralaryngeal and laryngeal activities (Laver, 2009).

Trewin, a veteran TV and radio presenter, stated that "[y]our voice will be your unique identifying trademark and a selling point. It will be recognized by everyone who hears you" (2013, p. 22). We believe that Trewin was referring to the effect of a media presenter using their voice and that her suggestions as listed in Table 2 centralise on how to use a certain voice property in order for one's voice to be "a selling point". Trewin's conception of the physical properties of voice not just includes volume (projecting), diction and intonation, but also incorporates what are traditionally known as delivery features like sounding calm or relaxed, breathing and pace. Emphasis in Trewin's sense is in essence intonation, as she advises that one should not be "formulaic" when placing stress on words (p. 35), which we think is synonymous with "monotonous intonation" in Collados Aís (1998), Zhang and Ke (2008), and Pöchhacker and Zwischenberger (2010).

Of special note is Trewin's relating voice to emotions in that one's voice needs to sound "calm and relaxing". This in our belief applies to interpreting, too, as interpreters can find themselves confronted with emotional situations, which will affect their voice (e.g. Harvey, 2003; Kurz, 2003; Valero-Garcés, 2005; Bontempo & Malcolm, 2011). Besson *et al.* (2005) regard emotional regulation, along with emotional intelligence, as "a sine qua non" of interpreters' skill kit. For the practical purpose of interpreter training, we propose that it would be useful to treat emotion as one of the physical properties of voice, which has to do with how one exhales air when vocalising, reflecting the dictionary meanings of voice (e.g., in Merriam-Webster Dictionary; Also see the Introduction section).

There is yet another important aspect concerning emotion for interpreters. This is interpreting emotions intended by a source speaker, as opposed to interpreters' displaying their own emotions as in our discussions above. As Wang and Huckvale (2018) have found, a pleasant interpreting voice displays controlled vividness of intonation. We accordingly propose that for an interpreter's voice to be heard as pleasant, the emotions intended by a source speaker should be moderated by an interpreter, as often observable in media interpreters[4].

As Table 2 (as well as Table 1) shows, there is not much literature on using voice to achieve a certain voice effect. Nonetheless, compared with the authors in Table 1, those in Table 2 who discussed or mentioned this provided more information in this regard by relating a certain voice

[4] It is not the scope of the current paper to define *pleasant voice* although this concept has often been referred to. We believe that with a clear definition of *voice* as we propose, the next step for interpreter training and interpreting assessment would be to clearly define *pleasant voice*. But this would be a different project for research.

effect to specific voice properties. For example, Schweda-Nicholson (1985) identified the misuse of volume and intonation as the causes for inadequate delivery. Trewin (2003) not merely provided a list of voice features for public speakers to watch out for, but also made specific suggestions on how to deploy these features for an effective voice. In the current paper, we follow a similar approach to conceptualising voice for interpreters.

Proposing a Definition of *Voice* for Interpreting

As we have shown, the existing literature on *voice* and *(un)pleasant voice* has not explicitly distinguished between the physical properties of voice on the one hand and the deployment of these physical properties to achieve a certain voice effect on the other (such as pleasant voice or unpleasant voice); relevant understandings of voice for interpreters tend to be anecdotal or inconsistent at times, and there appears to be a paucity of literature on what voice properties are responsible for interpreters' unpleasant voice or how to use/deploy voice properties to bring about a pleasant voice for interpreters. We think this is because the concept *voice* has not been defined in a formal manner, so much so that when (un)pleasant voice is the focus of discussion, it does not seem clear what voice properties lead to an interpreter's sounding (un)pleasant and how one needs to use/deploy their voice features in order to change from sounding unpleasant to sounding pleasant. This situation of elusiveness for the concept *voice* is similar to that of what the AIIC states about the concept *interpreting quality*, which the AIIC depicted as "that elusive something which everyone recognises but no one can successfully define" (1982). We feel the elusiveness surrounding the concept *voice* is not helpful for interpreting tutors and interpreting students, who both need pertinent feedback to discuss and improve interpreting performance.

In order for the concept *voice* to be a sufficiently useful didactic tool for both interpreting students and tutors/assessors, we propose to unify existing understandings of the concept. We propose that voice should be assessed as a non-verbal assessment criterion separated from *content* and *language* and should be regarded as a cluster of physical features with each feature defined. These features in the cluster together create a "voice image" that leads to an interpreter's voice to be heard as pleasant or unpleasant. It is these features that students and tutors could look at, both individually and collectively, when assessing voice quality in interpreting, and it is these features that could be commended when a voice is heard as pleasant or be trained when a voice is heard as unpleasant.

Our proposed conception of *voice* for interpreters corresponds to what is referred to as "the second factor" in Laver (2009, pp. 9-10; Also see the section "Elusiveness of *Voice* in Interpreting Literature"). The cluster of features of voice for interpreters as we propose is comprised of the following:

Table 3 a proposed definition of interpreters' voice

Interpreters' voice	**Description**
Intonation	Pitch, pitch movement and pitch contour
*Volume (*aka *projection)*	Degrees of loudness
Pace	pauses and delivery rate
Tone and emotional regulation	Air exhalation
Diction	Mouth shaping and phonation

Of note is that Lambeau (2006), Trewin (2013) and Horváth (2017) all also advise on the health of voice. We maintain this position, too, but consider it inappropriate to include this as part of voice quality assessment when a student/interpreter falls ill, unless they sound ill when they are not. If the latter is the case, then this should be assessed under *tone and emotional regulation*.

Our proposal is obviously built on the existing literature that has brought to light the importance of pleasant voice for interpreters, such as Bühler (1986), Kurz (1993), Kurz and Pöchhacker (1995), Moser (1996), Collados Aís (1998), Zhang and Ke (2008), and Pöchhacker and Zwischenberger (2010). Our proposed concept *voice* is an umbrella term, thus reflects De Gregoris's approach (2016). We have also borrowed the term "voice image" from Horváth (2017), having narrowed down its sense. Our proposed voice model is an attempt to formally define *voice* and provide non-overlapping and clear-cut subcategories so as to assess interpreters' voice for the effect pleasant voice in a comprehensive manner. This in our belief would potentially bring some new benefits for interpreter education and interpreting studies.

Firstly, voice is defined in terms of a set of mutually exclusive sub-categories. Such being the case, when assessing voice in terms of its pleasantness, both tutors and students could be clearer as to what to look at and what to advise on. For instance, we think Su's (2019) student participants would particularly benefit from the clear-cut cluster of voice properties when they assess each other's voice in their self-study. Our proposed definition would also potentially

help avoid double- or multi-penalties. Thus, for those assessing simultaneous interpreters in De Gregoris (2016), for example, they would find it easier to avoid penalising interpreters for their syntax when assessing the latter's voice. As a pleasant voice is the favoured effect of an interpreter using their voice, both tutors and students may now be able to pinpoint problematic voice aspects that lead to a voice being heard as unpleasant. As it could be one or more of the voice aspects listed in Table 3 that bring about an unpleasant voice, assessors would be able to identify which aspect or which aspects are responsible for an individual student's unpleasant voice. Likewise, if an interpreter's voice is heard as pleasant, tutors and students would also be able to identify specific voice element(s) that can be learnt from. Understandably, for individual interpreters or students whose voice is heard or judged as unpleasant, the factor(s) causing their voice to be so heard or judged can be different and subsequently remedial actions will be different, too. Table 4 presents a summary of some voice problems identified in the existing literature by applying our voice model.

Table 4 common causes for interpreters' unpleasant voice

Interpreters' voice	**Improper use of voice properties**
Intonation	Too high pitch, too drastic or monotonous pitch movement
Volume (projection)	Too loud, too weak
Pace	Disfluency, too fast, too slow
Tone and emotional regulation	Nervous, aggressive, critical, disappointed, anxious, harsh, indifferent, over-enthusiastic, over-friendly, personal
Diction	Mumbling, unintelligible sounds

As Fig.1 shows (See the "Introduction" section), the effect of a voice being pleasant or unpleasant has a bearing on the way of one using or deploying the physical properties of their voice. Our proposed definition of the *voice* concept would thus help make it clear that voice training for interpreters for the purpose of improving their voice quality is essentially training them to improve or correct the way of their using their voice properties related to an identified problem. Further along this line, our proposed definition of *voice* may have implications for developing self-learning tools for interpreting students so that they can use the tools to monitor their interpreting voice.

Secondly, our proposed concept *voice* would potentially help to define more clearly other non-verbal criteria used in interpreting assessment. In relevant literature, e.g., Su (2019) and Liu,

Chang and Wu (2008), *voice* is subsumed under *presentation*, which in Su (ibid., p. 179) consists of "voice, fluency, accent and other presentation features (e.g., microphone use, volume control)". With our proposed model, voice could be assessed as an independent rubric instead of being subsumed under *presentation*. The purpose of this new differentiation is to highlight the importance of voice training for interpreters, thanks to the existing literature that has shown the importance of a pleasant voice for interpreters. Our approach would generate two generic criteria for *non-verbal performance*: one for *voice*, and the other for *presentation*. To distinguish from voice performance, which relates to how interpreters use their vocal track and vocal cord, an interpreter's presentation is related to non-verbal performance that does not have to do with voice. This includes body posture, body gestures, facial expressions and eye movement. Presentation in this sense is assumedly what Bühler (1986) refers to as poise. Our approach would then potentially bring about a balanced assessment rubric where an interpreter's *interpreting performance* would be examined in terms of two generic categories: *verbal* and *non-verbal* as shown in Fig. 2, with the formal being assessed in terms *content* and *language*, and the latter in terms of *voice* and *presentation*, with each being able to be broken down further in an assessment sheet.

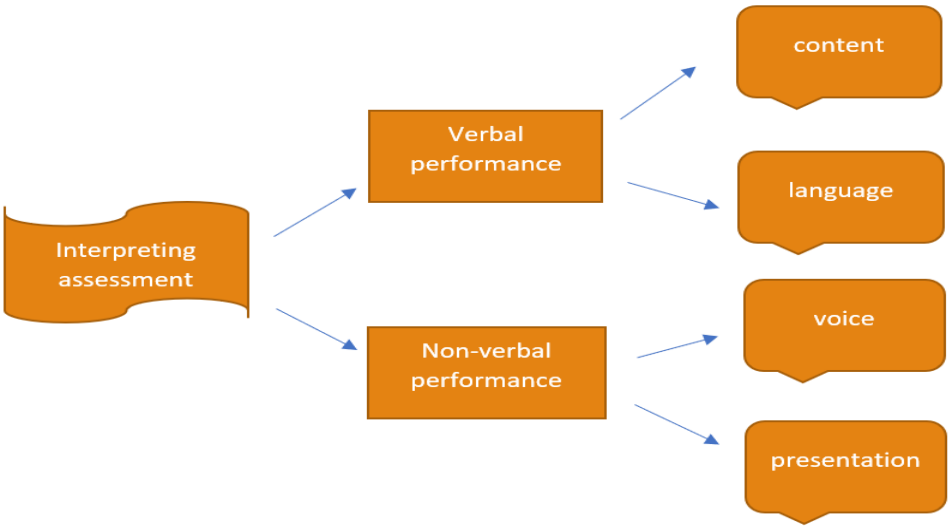

Fig. 2 Interpreting assessment criteria

It is worth reiterating that our proposed voice model has incorporated what has been traditionally known as delivery features (See Table 3). Our rationale is that at its core, delivery

has to do with voice production and thus it is appropriate to incorporate delivery features into voice.

Thirdly, our proposed concept of *voice* would potentially contribute to rigidity and comparability of research on voice and (un)pleasant voice for interpreters. As a pleasant voice is desired for interpreters' voice quality, it follows that with a clear set of defined parameters, our proposed definition of voice would make it easier to collect data for surveys and/or observations on voice or pleasant voice, thus helping interpreting scholars make conclusions in these regards and compare each other's research results. With our proposed definition, it would also be possible to survey the relative importance of the different voice properties to the ears of interpreters and interpreting service users, so as to inform interpreter education. Building on Collados Aís (1998), Zhang and Ke (2008), and Pöchhacker and Zwischenberger (2010), for example, one could hypothesise that intonation is a major voice factor contributing to the effect *pleasant voice* and then set out to collect data from interpreters or interpreting service users regarding their views. Additionally, as scholars such as Kurz (2003) indicate, it is important to study interpreting service users' view on interpreting quality so as to inform interpreter education. It is reasonable to assume that without a formal definition, when interpreting service users are surveyed for their opinion towards interpreters' voice, data on their perceptions may lack consistency, as evident in Iglesias (2013; Table 1). With our proposed definition of *voice*, data from future studies surveying interpreting service users' or interpreters' view towards interpreters' voice would turn out to be more comparable.

Conclusions

Horváth stated the following to highlight the importance of voice for interpreters and the importance of conscious voice training for the *pleasant voice* effect (2017, p. 233):

> Interpreters rarely consider their speech behaviour unnatural since they are rarely aware of it. One of the reasons for this lack of awareness can be found in lack of training at least in the case of those who have not received any formal training to become interpreters. However, voice consciousness, i.e. knowing one's voice and using it in a conscious way is one of the characteristics of professionalism in interpreting.

We hope that our proposed model for voice will provide a way for both interpreter trainers and interpreting students to better monitor the latter's voice in interpreting and thereafter further enhance voice consciousness.

While calling for voice consciousness, Horváth (ibid., p. 235) also pointed to some future research directions. Among the new research questions she raised, two are particularly relevant to our present paper. One is "Is voice consciousness enough to provide the interpreter with an increased control over the suprasegmental features of the speech they are producing?" The other is "To what extent is it possible for interpreters to prevent their emotions from being heard through their voices?" We call for research to employ our voice model and find answers to these two questions to further contribute to interpreter education and interpreting assessment. There could also be a third direction of inquiry. This is to survey the opinions of interpreting tutors and students, as well as interpreting service users, regarding the usefulness of our proposed voice model. Investigations in this regard would help us revise our model if/when necessary, in order for it to be optimally useful.

References

AIIC (the International Association of Conference Interpreters) (1982). Practical guide for professional interpreters. Geneva: International Association of Conference Interpreters.

AIIC (1999). Practical guide for professional conference interpreters. Accessed from: http://aiic.net/p/628, May 2020.

Amato, A. & Mack G. (2011). Interpreting the Oscar Night on Italian TV: an interpreter's nightmare? *The Interpreters' Newsletter*, 16, 37–60.

Alvstad, C., Greenall, A. K., Jansen, H. & Taivalkoski-Shilov, K. (2017). (ed.s) Textual and Contextual Voices of Translation. John Benjamins.

Besson, C., Graf, D., Hartung, I., Kropfhäusser, B. & Voisard, S. (2005). The importance of non-verbal communication in professional interpretation. Access from: http://aiic.net/p/1662, May 2020.

Bontempo, K. & Malcolm, K. (2011). An ounce of prevention is worth a pound of cure: Educating interpreters about the risk of vicarious trauma in healthcare settings. In Malcolm, K & Swabey, L (Eds). *In our hands: Educating healthcare interpreters* (pp. 105–130). Washington DC: Gallaudet University Press.

Bühler, H. (1986). Linguistic (Semantic) and Extra-linguistic (Pragmatic) Criteria for the Evaluation of Conference Interpretation and Interpreters. *Multilingua* 5(4), 231–235.

Cecot, M. (2001). Pauses in Simultaneous Interpretation: A Contrastive Analysis of Professional Interpreters' Performances. *The Interpreters' Newsletter*, 11, 63–85.

Collados Aís, A. (1998). Quality Assessment in Simultaneous Interpreting: the Importance of nonversbal Communication. Translated by Pamela Faber, *in (eds.)* Pöchhacker, F. & Shlesinger. M. *The Interpreting Studies Reader* (pp. 326–336). Routledge: Taylor & Francis Group. London and New York.

Chiaro, D., & Nocella, G. (2004). Interpreters' perception of linguistic and non-linguistic factors affecting quality: A survey through the World Wide Web. *Meta*, 49(2), 278–293.

De Gregoris, G. (2016). A proposal for an assessment of voice quality in TV broadcast simultaneous interpreting through a gestaltic approach: theoretical paradigm for a new questionnaire. *MonTI* Special Issue 3, 273–306.

Gilles, A. (2013). *Conference Interpreting: a Student's Practice Book*. Routledge: Taylor & Francis Group. London and New York.

Harris, B. (2015). The interpreter's voice. *The savvy Newcomer: Tips and resources for translators and interpreters*. Accessed at: https://atasavvynewcomer.org/2015/05/19/interpreters-voices-1/, May 2020.

Harvey, M. (2003). Shielding yourself from the perils of empathy: The case of sign language interpreters. *Journal of Deaf Studies and Deaf Education*, 8, 207–213.

Hermans, T. (1996). The Translator's Voice in Translated Narrative. *Target*, 8 (1), 23–48.

Horváth, I. (2017). The speech behaviour of interpreters. In: *Across Languages and Cultures*, 18(2), 219–236.

Hubscher-Davidson, S. (2009). Personal Diversity and Diverse Personalities in Translation: A Study of Individual Differences. *Perspectives*, 3, 175–192.

Iglesias F. E. (2013). Understanding variability in interpreting quality assessment: User's sex and judgments for pleasant voice. In Way C., Vandepitte S., Meylaerts R. & Bartłomiejczyk M. (ed.s) *Tracks and Treks in Translation* Studies (pp. 103–125). John Benjamnins.

Kurz, I. (1993). Conference interpretation: Expectations of different user groups. *The Interpreters' Newsletter*, 3, 13–21.

Kurz, I. (2003). Physiological stress during simultaneous interpreting: a comparison of experts and novices. The Interpreters' Newsletter, 12. EUT - Edizioni Università di Trieste, 51–67.

Kurz, I. & Pöchhacker F. (1995). Quality in TV Interpreting. *Traslatio-Nouvelles de la FIT- FIT Newsletter* 14:3-4, 350–358.

Lambeau, M. (2006). Developing the interpreter's voice. In: *The ATA Chronicle*, 26–27.

Laver, J. (2009). *The Phonetic Description of Voice Quality*. Cambridge University Press.

Liu, M., Chang, C. & Wu, S. (2008). Interpretation Evaluation Practices: Comparison of Eleven Schools in Taiwan, China, Britain, and the USA. *Compilation and Translation Review*, 1(1), 1–42.

Lu, X. (2007). Interpreting Quality: Basic Criteria and Multiple Perspectives. *Language and Translation*, 4, 40–45.

Ma, N. (2010). 课堂口译质量评估的必要性及评估体系的建立[The assessment model for in-class interpreting assessment]. *Journal of Jiamusi Education Institute*, 1, 157–158.

Merriam-webster Dictionary: https://www.merriam-webster.com/dictionary/voice (Accessed August 2020)

Moser, P. (1996). Expectations of users of conference interpretation. *Interpreting*, 1(2), 145–178.

Nolan, J. (2012). *Interpretation: Techniques and Exercises*. Bristol: Multilingual Matters.

Perng, H. [彭輝榮] (2006). Chinese-English Interpreter Training Procedure and Text Exercise Sample Designs [中英口譯訓練過程與語篇練習範本設計]. *Spectrum: NCUE Studies in Language, Literature, Translation, and Interpretations*, 1, 185–203.

Pöchacker, F. & Zwischenberger C. (2010). Survey on quality and role: conference interpreters' expectations and self-perceptions. Accessed from https://aiic.net/page/3405/survey-on-quality-and-role-conference-interpreters-expectations-and-self-perceptions/lang/120, May 2020.

Pöchhacker, F. (2011). Researching TV interpreting: selected studies of US presidential material. In: *The Interpreters' Newsletter*, 16, 21–36.

Russo, M. (2005). Simultaneous film interpreting and users' feedback. *Interpreting*, 7(1), 1–26.

Schweda-Nicholson, N. (1985). Consecutive Interpretation Training: Videotapes in the classroom. *Meta*, 30(2), 148–154.

Trewin, J. (2013). *Presenting on TV and Radio: An insider's guide*. Hoboken. NJ: Taylor and Francis.

Wang, C. & Huckvale, M. (2018). Elucidate the elusive concept of 'pleasant voice' for interpreters: a case study. *2nd Conference of the World Interpreting and Translation Teaching Association (WITTA)*. University of Westminster, London 13 - 14 Oct 2018.

Wang, W. & Mu, L. (2013). "期待产品": 口译产品研究中的一个模型. *Foreign languages and their teaching*, 5, 73–77.

Zhang, W. & Ke, F. (2008). The Effect of Interpreting Users in Interpreting Quality Assessment [从口译用户看口译质量评估]. *Foreign language research* [外语学刊], *3*, 114–118.

Valero-Garcés, C. (2005). Emotional and psychological effects on interpreters in public services --- A critical factor to bear in mind. *Translation Journal*, 9(3). Accessed from https://translationjournal.net/journal/33ips.htm, May 2020.

Su, W. (2019). Interpreting quality as evaluated by peer students. *The Interpreter and Translator Trainer*, 13 (2), 177–189.