Discussing the Limits of Objective Audio Description
(University of Hildesheim, 6 October 2022)

Neutrality in audio description delivery

Anna Matamala

Anna.matamala@uab.cat

TransMedia Catalonia

Departament de Traducció, d'Interpretació i d'Estudis de l'Àsia Oriental

Universitat Autònoma de Barcelona

UAB
Universitat Autònoma
de Barcelona

TRANS
MEDIA
CATALONIA

Audio description

WHAT	HOW
Content selection (discussion about objectivity/neutrality)	Phrasing Voicing

Voicing

- Voicing: uniform, adapted, emphatic (Cabeza-Cáceres, 2013)

- Describers have been traditionally "encouraged to use a particularly **neutral** way of speaking" and "a neutral delivery has come to be recognized as the norm" (Fryer 2016, 88)

- UNE 15320: "voicing must be **neutral** with appropriate intonation"

Research

What is "a neutral voice"?

- Research with María Machuca and Antonio Ríos, phoneticians (see publications in the reference list).

- Contributions from Inés Fernández (MA Thesis).

- Part of the NEA and RAD projects.

- Thanks to all participants and to associations such as ACPA or ACB reaching out to participants.

Neutrality in prosody

Neutral = unmarked (Crystal 2008)

Non-emotional (Busso, Lee and Narayan 2007)

Emotions: specific prosodic features related to
 a) loudness (intensity, amplitude, volume)
 b) rate
 c) pitch (F0 value, type of voice)

Still, underresearched topic.

Many questions related to AD

- What prosodic features present the voices used in ADs?

- What voices are selected by users as neutral?

- What voices are selected by users as pleasant?

- Do neutral and pleasant voices correlate?

Many questions related to AD

- Are there any differences across languages?

- How do users define a "neutral voice"?

- Are there differences between those familiar with AD and those not familiar with AD?

- Between persons with and without sight loss?

Research by Machuca, Ríos & Matamala

- Analysing a corpus of AD: acoustic analysis.

- Perception test with users with and without sight loss. They were also asked to define a "neutral voice".

- **Focus today: definition (as part of the perception test).**

Selection of the voices

- Based on the acoustic analysis.

- Visuals into Words corpus (webs.uab.cat/easit).

- 10 AD in Catalan + 10 in English + 10 in Spanish (50% male, 50% female).

Selection of the voices: acoustic analysis

- AD units divided into intonation groups and analysis with PRAAT of:

    - F0 value (pitch), in Hz.
    - Average amplitude, in dB.
    - Duration, in ms.

- Selection of 3 male voices and 3 female voices per language.

Stimuli

- 18 stimuli (9 contrasting female voices, 9 male voices).

- No interaction with the AD (need to start with shorter stimuli).

Participants

- Spanish: 60 (31 with sight loss, 29 without).

- Catalan: 81 (35/46).

- English: 71 (40/31).

Total: 212 participants.

Procedure: online form

- Welcome and consent.
- Sex / person with or without sight loss.

1. Please define what a neutral voice is for you.
2. Click on different pair of voices and indicate which sounded more neutral.
3. Click on different pair of voices and indicate which was more pleasant.

Focus today: 1 (and 2).

Results: voice definition

- 206 replies from:
  - 98 participants with sight loss
  - 108 without sight loss
  - 139 women
  - 67 men

- Thematic analysis.

Themes

- Emotions
- Pitch
- Inflection
- Accent
- Articulation, pronunciation
- Volume
- Gender
- Synthetic/natural voices.
- Style
- AD-related aspects

Themes: emotion

- Emotions (34.46%)
  - "a voice lacking in emotion"
  - "a voice not expressing feelings"
  - "a voice that does not provoke any specific feeling in the listener"
  - " a neutral voice is a voice with little expression, but not so much that it determines my reactions or emotions"
  - "not too excited"
  - "one that does not show excessive emotion"
  - "with minimal emotion"
  - "neither too expressive nor too monotone"
  - "a voice that is less involved in the film being described"
  - "the reader is not engaged in what they are reading, not responding or reflecting the content of what they are reading"

Themes: emotion (AD-related)

- Some evaluative comments from participants with sight loss:

  - "the voice must sound different enough from the other voices involved in the movie so that the describer's voice is not confused with the voices of the characters. It can show emotion."
  - "neutral voice in audio description should contribute to scene understanding, not guiding you in the emotional interpretation."

Themes: pitch and inflection

- Pitch: 51 participants (24.75%).
- Inflection (changes in tone): 48 (23.3%).
  - "Neither too high in pitch nor too low".
  - "Mid-range".
  - "Medium in tone"
  - "Not too high".
  - A neutral voice is "low" (EN, one participant).
  - "Minimal modulations in tone, not overly high- or low-pitched"
  - It "does not have inflection" and "is consistent in tone".
  - "It can have some inflectiu but not a dynamic amount".
  - "With little inflection and range".
  - "a neutral voice is one that doesn't change much in pitch or tone, but also isn't so flat as to sound monotone."

Themes: pitch and inflection (AD-related)

- "With enough inflection to hold the listener's attention, but not so much that the listener is distracted from the plot"

- "with some vocal modulation to keep it interesting, not robotic".

- "a neutral voice is one that displays neither exaggerated nor reduced intonation and stress patterns, though I acknowledge that what constitutes 'normal', 'exaggerated', and 'reduced' is somewhat subjective."

- "with proper but not too much inflection. Without good inflection, describers can ruin a show."

Themes: accent

- Accent: 30 participants (14.56%)
    - "no accent"
    - "no dialectal traces"
    - "free from regionalism accent"
    - "without a noticeable accent"
    - "without heavy accents"
    - "without a pronounced regional accent".

    - EN: "A mid-Atlantic American accent", "a neutral American accent"
    - The origin of the speaker should not be identified through the accent.

Themes: articulation, pronunciation

- Good articulation and clear pronunciation: 16 (7.76%)
  - "clear delivery with crisp articulation"
  - "clear and enunciated"
  - "good diction"
  - "with good pronunciation and vocalization"
  - "a correct pronunciation"

- Understandable: 18 (7.36%)
  - "clear words"
  - "it is not hard to figure out what the person is saying"
  - "easily understood"
  - Focus on content: "One that explains things in detail clearly".

Themes: volume & speed rate

- 13 participants (6.31%)

    - "moderate pace and loudness"
    - "speaks at a medium tempo"
    - "no whispering, no yelling"
    - "not going to the extremes of speed"
    - "without any significant change in volume level"
    - "calm"
    - "soft"
    - "steady, not too fast, not too slow"

Themes: gender

- 12 participants (5.82%)

  - "neither male or female"

  - "genderless"

  - "it can be male or female", as this aspect is not relevant.

  - Only one American female writes down that a neutral voice is "female".

Themes: synthetic/human voices

- A natural voice is
    - "human-like"
    - "does not sound fake"
    - "not too mechanical – I can listen to a computer screen reader for that"
    - "sounds natural"
    - "does not sound like a machine".

Confusion in the study purpose among the American participants and interesting feedback.

Themes: AD features

- 18 participants (8.73%)
  - "not distracting from the media being described"
  - "a voice which gives needed visual information without distracting from the other auditory information being given"
  - "Something that can easily blend into the background as a stream of information to the listener, rather than distracting them with distinct stressing or making themselves the star of the show"
  - "one that blends in and doesn't distract from the context it is in".
  - "one that does not lean towards positive or negative and does not influence my assessment of what is being stated".
  - a neutral voice "must sound different enough from the other voices involved in the movies so that the describer's voice is not confused with the voices of the characters"

Additional AD-related comments

- "not sure neutral is a good thing in an audio describer"
- "one that describes the action and environment without interjecting their own opinions or perception"
- "something not expressing an opinion"
- a neutral voice "doesn't just read the script but supports the story. Thus there is no one neutral voice"
- "It does not convey a subjective interpretation but still is consonant with the action/material being described"
- "it helps you understand the scenes, not guide you in the emotional interpretation"
- "without connotations"
- "without evaluations"
- "without totally lacking feelings, describes the scene objectively"
- a neutral voice should be "objective."

A step further

- MA thesis by Inés Fernández (2021)

- Perception test focusing on pitch in Spanish:
  - when does a neutral voice cease to be neutral?
  - Please define what a "neutral voice" is.

- 26 participants with AD knowledge + 29 without.

<div align="right">

Fernández, Machuca and Matamala (forth.)

</div>

Responses

- Participants defining only 1 features:
  - 82.14% without AD knowledge.
  - 34.86% with AD knowledge.

- Similar responses except for:
  - Prioritisation.
  - Reference to "objectivity" only from those with AD knowledge.

Themes: similar to the previous ones

60% of the responses focus on:

- Emotions ("it does not generate feelings", "it does not transmit an emotion")

- Accent ("neutral or standard accent")

- Pitch ("not too low, not too high")

- Inflection ("flat", "monotone", "without variation")

Other themes

- Objectivity (only mentioned by participants with AD knowledge): "voice that transmits the message in the most objective possible tone", "a voice transmitting objectivity".

- Speed: "Paused", "calm".

- Other views related to musicality, timbre, among other features.

Some thoughts

Impact of the objectivity/subjectivity or neutrality debate on some of the definitions.

How do subjective definitions correlate with prosodic analysis?

Results: voice selection

		Spanish		English		Catalan	
		With sight loss	Without	With sight loss	Without	With sight loss	Without
		More neutral					
Male	Pitch	med		+	-	+	
	Rate	-		+	-	-	
	Amplitude	-		+	-	-	
Female	Pitch	-		-	med	+	
	Rate	-		+	med	-	
	Amplitude	med		+	-	med	
		Less neutral					
Male	Pitch	+		-		-	
	Rate	+		+		med	
	Amplitude	+		med		med	
Female	Pitch	+		+		med	
	Rate	+		-		+	
	Amplitude	+		+		+	

Machuca & Matamala, forthcoming, Babel

TRANS MEDIA CATALONIA

UAB Universitat Autònoma de Barcelona

Conclusions

Impact of the objectivity/subjectivity or neutrality debate on some of the definitions.

Need to define a "neutral voice" from a prosodic point of view.

Need to do research at different levels (isolated stimuli, AD in context, etc.).

And you?

How would you define a "neutral voice"?

Would that definition be different if I ask you "How would you define a "neutral voice in AD"?

And what if I ask: "Should the voice of the audio describer be neutral?"

And if your reply is "it depends", then I ask: "On what does it depend?".

Discussing the Limits of Objective Audio Description
(University of Hildesheim, 6 October 2022)

Neutrality
in audio description delivery

Anna Matamala

Anna.matamala@uab.cat

TransMedia Catalonia

Departament de Traducció, d'Interpretació i d'Estudis de l'Àsia Oriental

Universitat Autònoma de Barcelona

UAB
Universitat Autònoma
de Barcelona

TRANS
MEDIA
CATALONIA