Audio description research: what next?

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

Transmedia Catalonia

https://grupsderecerca.uab.cat/transmedia/

AD research in Transmedia Catalonia

ALST

ALST

"Linguistic and sensorial accessibility: techologies for voice-over and audio description", 2013-2015 (Spanish funding).

PhD thesis by Anna Fernández-Torné on AD and technologies.

Matamala, Anna (2016) "The ALST project: technologies for audio description". In Matamala, Anna; Orero, Pilar (eds) *Researching Audio Description. New Approaches*. London: Palgrave Macmillan, 269-284.

Rationale

Speech recognition

Machine translation

Text-to-speech

Voice-over? AD?

Speech recognition

Soundtrack extraction, speech activity detection, speaker diarisation and speech-to-text transcription.

Film "Closer" (English and Catalan)

DER & WER measures: low performance.

Source: Delgado, Matamala and Serrano (2015)

Machine translation

Main experiment: 12 AVT MA students

Human creation/Human translation/post-editing (EN>CAT)

Temporal, technical and cognitive effort (InputLog)
Perceived effort (questionnaire)

Source: Fernández-Torné & Matamala (2015)

MT in AD: results

- Temporal effort
- Technical effort
- Cognitive effort (pause-to-word ratio)

- Subjective measures

MT: speech synthesis

Pretest: select the "best" human/natural male/female voice

Main experiment with 67 persons with sight loss.

Main results:
- Natural voices, higher values.
- Overall impression/acceptance: mean higher than 3.2
- 94% TTS acceptable solution, not preferred one.

Source: Fernández-Torné & Matamala (2015)

In line with previous research by Szarkowska *et al.* in Poland

What next?

What about other languages?

- ALST: Spanish<>Catalan (Matamala & Ortiz Boix 2016)
- TRICS project on MT and AD (University of Antwerp), PhD by Kim Steyaert

What about automatic translation of images?

- Image captioning/automatic video description: MeMAD project (Braun, starr, Delfani; Koponen, Hirvonen).

VIW

https://pagines.uab.cat/viw/

The Corpus

The VIW corpus is made up of audio descriptions in different languages of the same short film, and it incorporate a **search engine** (forthcoming). It currently includes 10 professional audio descriptions in English, 10 professional audio descriptions in Spanish and 10 professional audio descriptions in Catalan. All 30 audio descriptions were commissioned to professionals, who were remunerated. It also includes 10 audio descriptions in Spanish and 7 audio descriptions in Catalan made by students (text only, not recorded audio/video files). We hope to expand it in the near future with more audio descriptions. You are invited to read the section "Contribute".

Audio and video files with audio description by professional describers

▶ English   ▶ Catalan   ▶ Spanish

Text files of audio descriptions by students

▶ Catalan   ▶ Spanish

ACCESS TO THE CORPUS SEARCH ENGINE

All materials are also available at UAB's open access repository: https://ddd.uab.cat/record/147267.

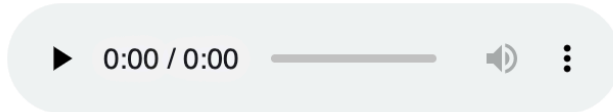TRANS MEDIA CATALONIA   UAB Universitat Autònoma de Barcelona

ENGLISH AUDIO DESCRIPTIONS

▶ English audio descriptions   ▶ Catalan audio descriptions   ▶ Spanish audio descriptions

All materials are also available at UAB open access repository: https://ddd.uab.cat/record/147267.

Able (NZ):

▶ 0:00 / 0:00 🔊 ⋮

▶ ▭▭▭▭▭ 0:00 🔇

Bridge Multimedia (USA):

▶ 0:00 / 0:00 🔊 ⋮

▶ ▭▭▭▭▭ 0:00 🔇

BTI Studios (UK):

▶ 0:00 / 0:00 🔊 ⋮

List of available corpora

You can also use the [BROWSE facility](#) to look for a specific file.

- **WHW-EN-Pr**: The WHW-EN-Pr corpus contains professional audio descriptions in English for the short movie "What Happens While".
[access](#)

- **WHW-ES-Pr**: The WHW-ES-Pr corpus contains professional audio descriptions in Spanish for the short movie "What Happens While".
[access](#)

- **WHW-CA-Pr**: The WHW-CA-Pr corpus contains professional audio descriptions in Catalan for the short movie "What Happens While".
[access](#)

- **WHW-ES-St**: The WHW-ES-St corpus contains audio descriptions in Spanish for the short movie "What Happens While" made by students. [access](#)

- **WHW-CA-St**: The WHW-CA-St corpus contains audio descriptions in Catalan for the short movie "What Happens While" made by students. [access](#)

---

https://transmediacatalonia.uab.cat/web/

Actions for corpus 'WHW-EN-Pr'

- **Filmic annotations:** you can get the .eaf file with all filmic annotations or the html version

---

- **Simple string search:** this allows simple string searches on the entire WHW-EN-Pr corpus access

---

- **AD unit, sentence and word countings:** this provides some figures about AD units, sentences and words in the WHW-EN-Pr corpus access

---

- **AD distribution in the timeline:** you can get a compact view or an expanded view of the ADs distribution in the timeline

---

- **Verb distribution in the timeline:** you can choose a verbal semantic class and see its distribution in the timeline. access

---

- **AD similarity:** see text similarity among files in the corpus access

---

- **Word frequency by part of speech:** you can see word frequency by part of speech tag access

---

- **Word frequency by provider:** this gives the number of verbs, nouns, adjectives and adverbs per provider. access

---

- **Verbs & semantic classes:** you can inspect verbal semantic classes access

---

- **Nouns & semantic classes:** you can inspect nominal semantic classes access

---

- **Adjectives & semantic classes:** you can inspect adjectival semantic classes access

---

- **Adverbs & semantic classes:** you can inspect adverbial semantic classes access

---

What next?

Corpus research in AVT:

- dubbing (Pavia, Forlì, American Movie Corpus/ Matamala, Romer—Fresco, Baños);

- subtitling (Veiga corpus, Scandinavian subtitling corpus, CORSUBIL, ESIST, SUMAT, etc.)

- AD (TIWO, TRACCE, Pear Tree Project, Reviers, MPII-MD by Rohrbach).

- Issues: open access, multilingual.

- The way forward: expanding it & analysing the data.

References

Matamala, Anna (2019) "The VIW corpus: multimodal corpus linguistics for audio description analysis". *RESLA. Revista Española de Lingüística Aplicada,* 32:2, 515-542.

Matamala, Anna (2018) "One short film, different audio descriptions. Analysing the language of audio descriptions created by students and professionals". *Onomazéin*, 41, 185-207.

NEA
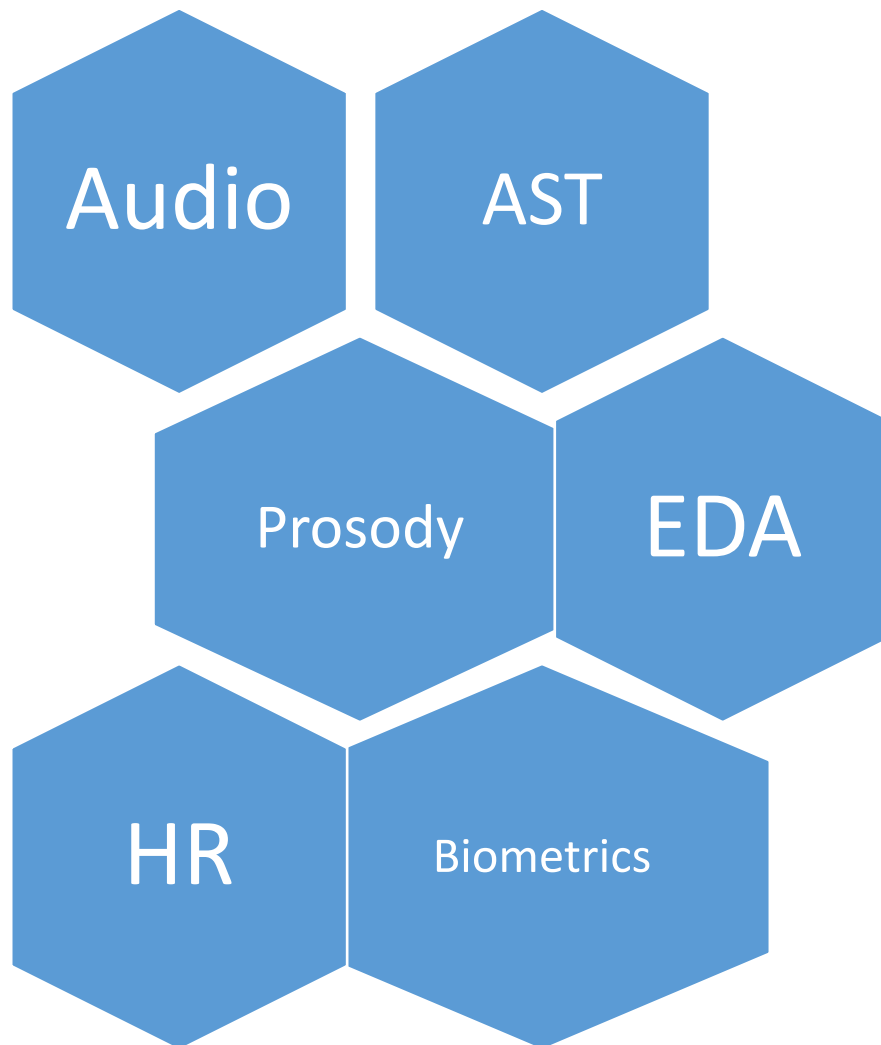
Sources:

Matamala et al (2020): "Electrodermal activity as a measure of emotions in media accessibility reserach"

Iturregui-Gallardo PhD on AST and emotional activation: voice-over versus dubbing effect. (*Perspectives*, 2020)

RAD
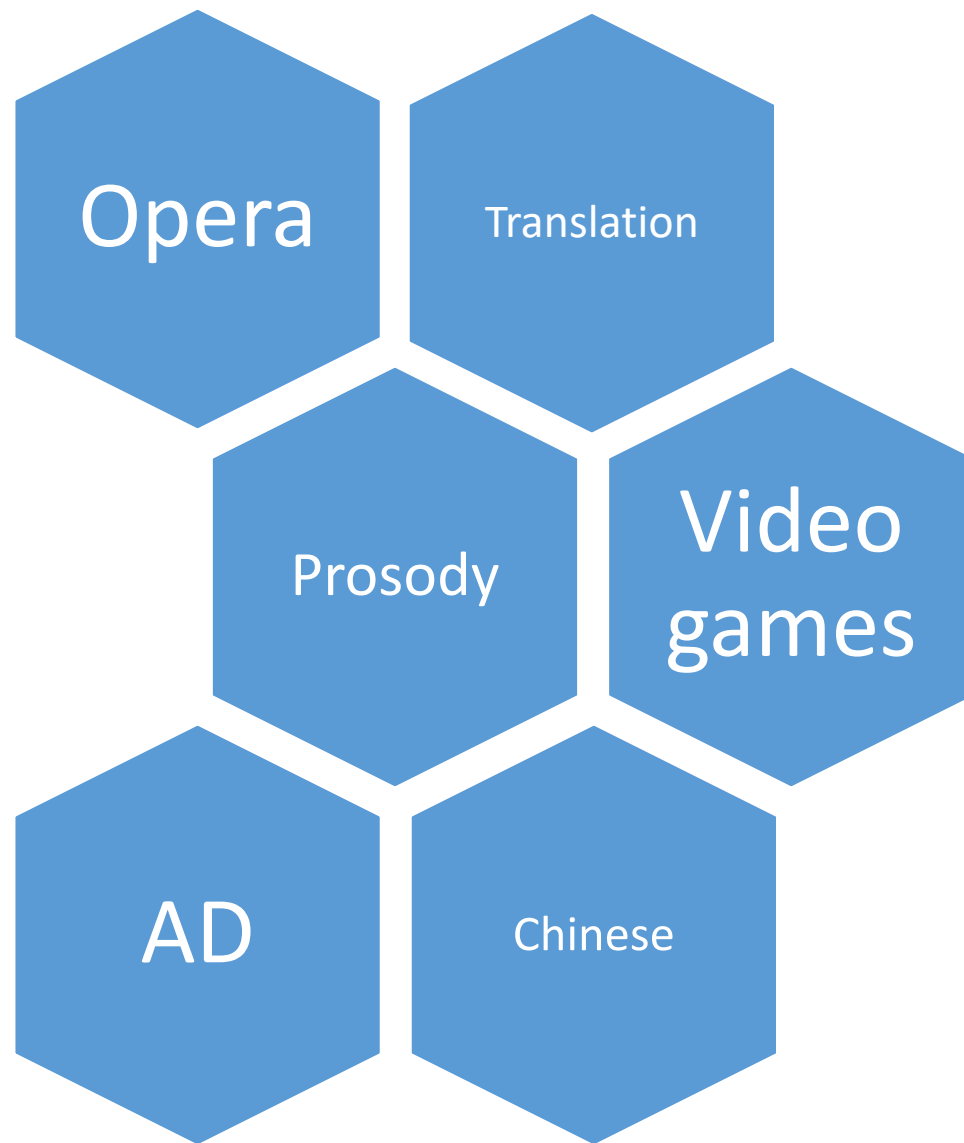
RAD: prosody

Voices in AD:

"a neutral delivery has come to be recognized as the norm" (Fryer 2016: 88)

"voicing must be neutral with appropriate intonation" (Spanish standard)

WHAT DOES IT MEAN?

Prosody

Descriptive study (using VIW corpus) in CAT, ES, EN:

- duration/speed
- intensity/volume
- pitch

Perception study:
- what is a neutral voice?
- select neutral/preferred voice (different pairs)

Results and ongoing work

- Non-neutrality, easier to identify than neutrality

- Differences across languages

  - Pitch
  - Volume
  - Speed

Results and ongoing work

- Qualitative comments:
  - Emotions, expressivity
  - Pitch
  - Volume
  - Speed
  - Changes in the previous parameters
  - Accent
  - Articulation, pronunciation
  - Unobtrusive
  - Gender

What next?

- Modifying values independently to better understand neutrality.

- Preferences versus objective data.

- Better understanding of audio/voice features (audio + description) and vocal skills.
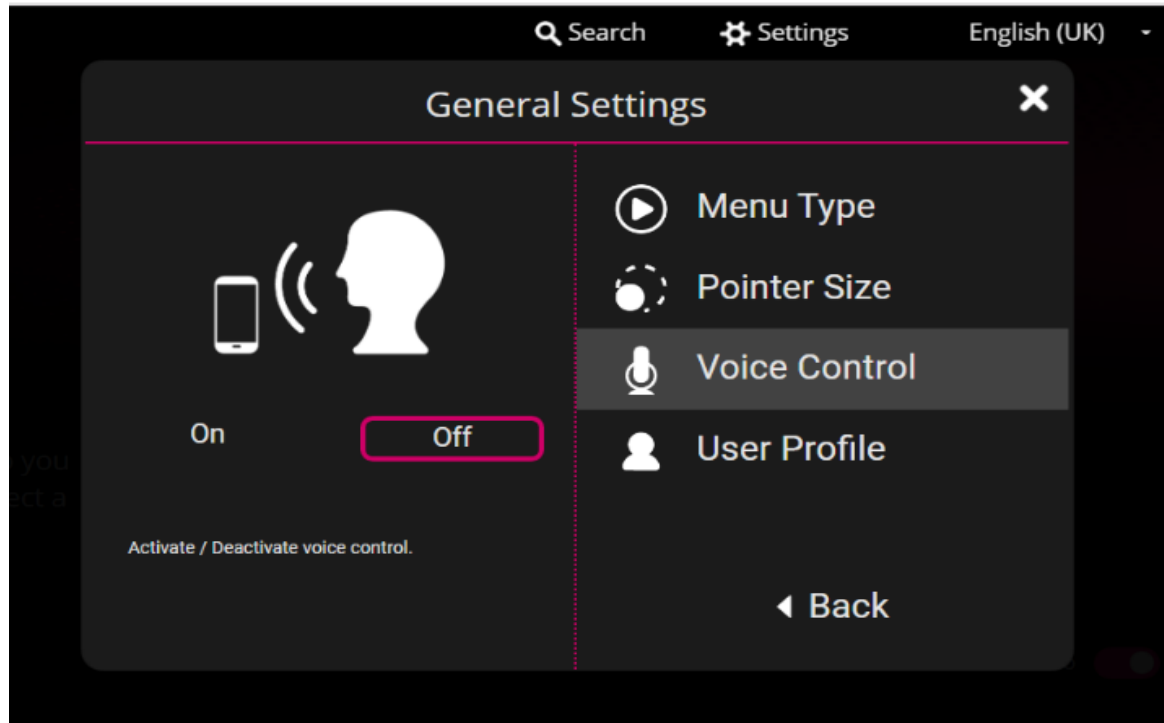
AD in 360º videos: services
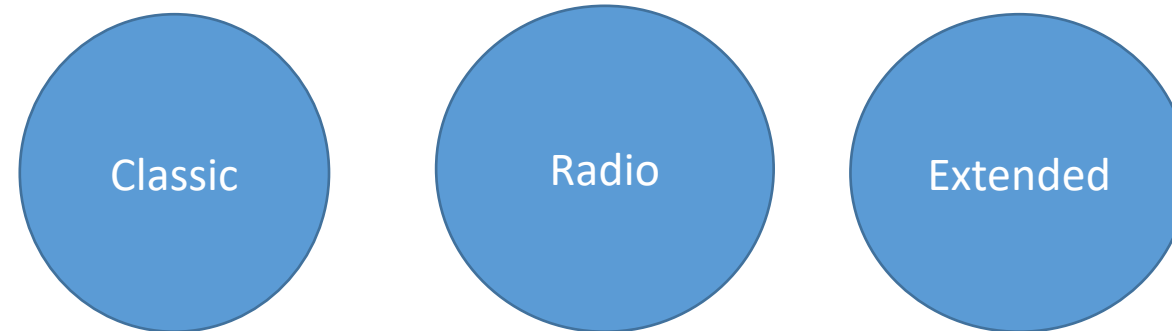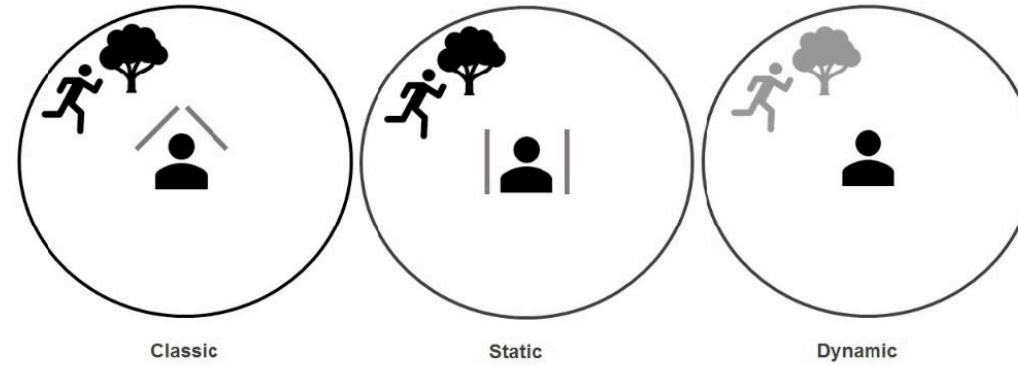
Sources: Phd by Anita Fidyka, Fidyka & Matamala (2018), D.5.4 IMAC deliverable

AD in 360º videos: editor
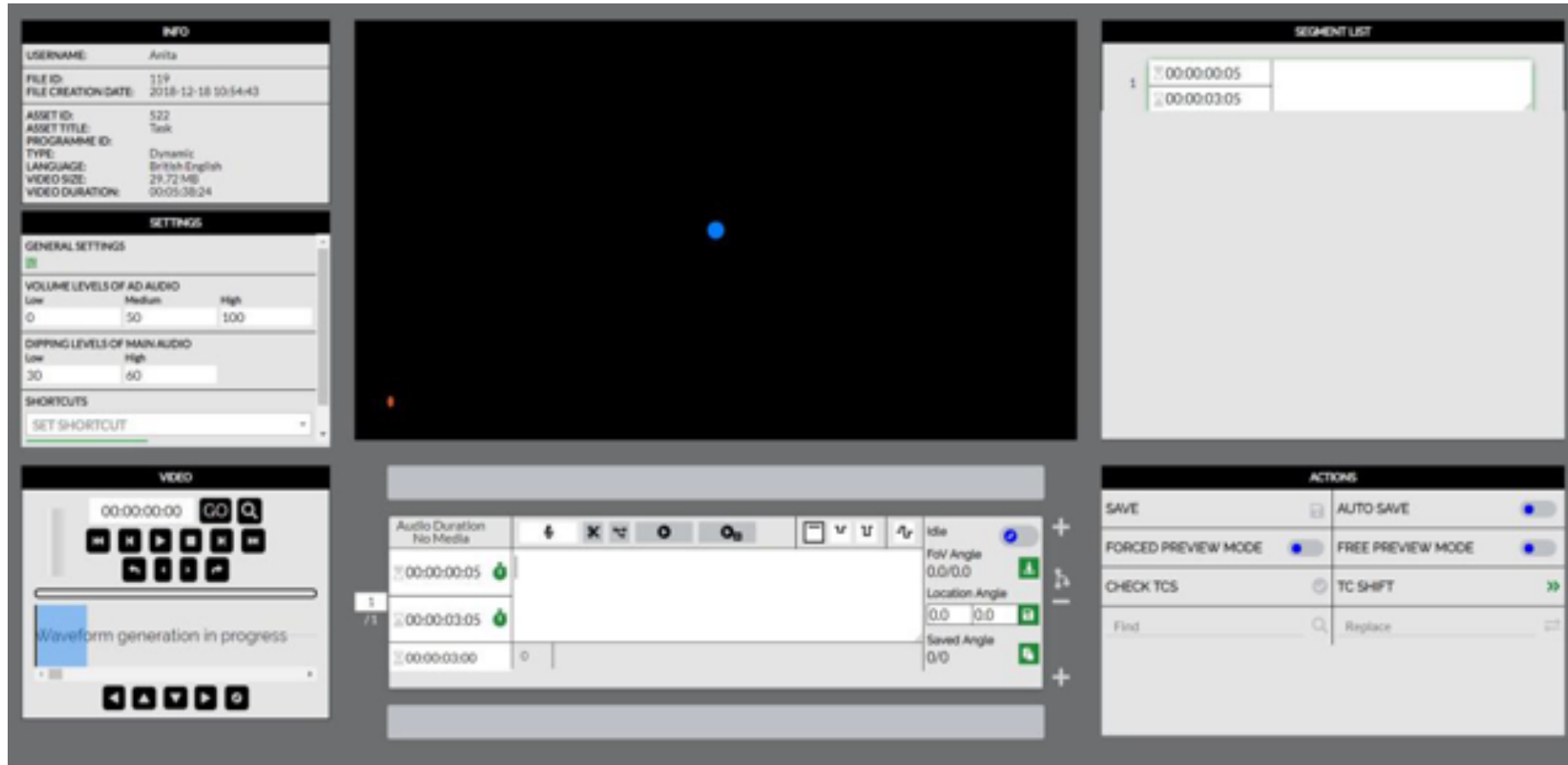
Source: Fidyka & Matamala (2019)

AD in 360º videos: player

What next?

- from 360º videos to interactive virtual reality
- from factual content to fictional content
- augmented reality
- videogames

EASIT

pagines.uab.cat/easit

Online Multiplier Event
June 2021. Free!

Erasmus + Strategic Partnerships in Higher Education.

2018-1-ES01-KA203-05275

September August 2021

✖ Before

Medicaid: Apply if you are aged (65 years old or older), blind, or disabled and have low income and few resources. Apply if you are terminally ill and want to receive hospice services. Apply if you are aged, blind, or disabled; live in a nursing home; and have low income and limited resources. Apply if you are aged, blind, or disabled and need nursing home care, but can stay at home with special community care services. Apply if you are eligible for Medicare and have low income and limited resources.

✔ After

You may apply for Medicaid if you are:

- Terminally ill and want hospice services
- Eligible for Medicare and have low income and limited resources
- 65 years old or older, blind, or disabled and have low income and few resources and:

  - Live in a nursing home
  - Need a nursing home care but can stay at home with special community care services

Source: https://www.plainlanguage.gov/examples/before-and-after/medicaid-eligibility/

ER

Top tips to feel better during the coronavirus pandemic

Easy read information

During this coronavirus pandemic, you may be feeling a lot of different emotions.

This is okay.

We all have different ways of coping.

These tips may work for different people at different times.

Can we create easy-to-understand

- subtitles?
- audio descriptions?
- audiovisual content (eg. audiovisual news)?

Is it feasible? Is it needed?

If so, how can we train these new professional profiles?

(Erasmus + > educational projects)

1. State of the art: training and practice.

2. How to transfer easy-to-understand language into the audiovisual world?

3. New professional profiles? What skills?

4. How to train them?

5. What educational content is needed?

6. How to certify training?

E2U AD focus group results:

- What if the dialogues are difficult

to follow? Differences between genres (audioguide versus film AD)?

- Is this really AD?
- How to cater for different user needs?
- Extended AD?
- Easy-to-listen?

AMÁS Fácil adapta los audios explicativos de la obra "La autora de las Meninas", que Teatro Accesible pone en escena el 14 de mayo en Barcelona

4 mayo, 2017 / en Actualidad / por admin grupoamas

"Ven a ver esta obra de teatro, la vas a entender".

https://www.grupoamas.org/amas-facil-adapta-los-audios-explicativos-de-la-obra-la-autora-de-las-meninas-que-teatro-accesible-pone-en-escena-el-14-de-mayo-en-barcelona/

Taylor and Perego (2021)

About 10 metres beyond the font are the Great West Doors. Made of deep brown oak, they're 9 metres high and about 4 metres wide, in a deep, arched alcove. Artwork is often exhibited in front. The doors are decorated by carved rows, alternating squares and rectangles, interspersed with semi-spheres, the size of an eating apple, small enough to cup in your hand. This grand entrance is reserved for when visitors such as HM The Queen attend the cathedral, while most visitors enter through the smaller doors on either side.

(VocalEyes, 2012)

The Great West Doors are beyond the font.
They are made of brown oak.
Oak is a type of hard durable wood.
They are very big.
They are 9 metres high
and 4 metres wide.
They form an arch at the top.
Often, artists display their work
in front of these doors.
Rows of squares and rectangles
decorate the doors.
Half spheres decorate the doors too.
The squares, rectangles and half spheres
are cut with care into the wood.
Only important visitors use these doors.
For example,
Her Majesty The Queen can use these doors.
Ordinary visitors use smaller doors.

What next?

More research is needed to expand on the works

by Orero & Bernabé-Caro,
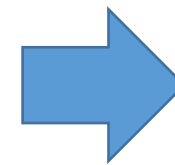by Elisa Perego and
by other EASIT team members.

What next?

What next?

Research on:

- Technology and workflows
- Less researched languages
- <u>Audio</u> + description
- Training
- Theoretical approaches (opening the lense)
- Multimodality
- Universal design/born accessible/accesible (filmmaking)

➡ **Knowledge transfer**

Sharing knowledge

Media Accessibility Platform

http://mapaccess.org

MAP media accessibility platform BETA

ABOUT | ACCESSOMETER | RESEARCH | TRAINING | MODALITIES | STAKEHOLDERS | EVENTS | NEWS

Welcome to MAP (beta version)

Welcome to MAP, the Media Accessibility Platform, a unified atlas charting the worldwide landscape of research, policies, training and practices in this field. MAP aims to make media accessible to all, regardless of sensorial and linguistic barriers.

Read more

PROJECTS

Understanding Media Accessibility Quality

Project duration **2017-2019**

Leading Institution **Autonomous University of Barcelona**
Funding bodies **European Commission**
Modalities **Media Accessibility**

LATEST NEWS

02/10/2018 **Languages and the Media programme online**
23/04/2018 **MAP as media partner of 12th Languages and The Media conference**
23/04/2018 **12th Languages and The Media Conference - Berlin 2018**
08/03/2018 **UMAQ conference: progamme online**
04/11/2017 **A new Collection of Movies with video-description**
19/10/2017 **EU one step closer to Accessibility Act**

October 2018

Su	Mo	Tu	We	Th	Fr	Sa
30	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	1	2	3

PUBLICATIONS

Audiodescrição Didática

Type **PhD thesis**   Year **2016**

Author(s) **Elton Luiz Vergara Nunes**
Modalities **Audio Description**

TRAINING COURSES

Cursos de Audiodescrição

Course Type **Professional training course**

Institution **Azimuth e-Learning**   Year **2017-2018**
Modalities **Audio Description**

EVENTS

XXXVIII Semana do Tradutor

Type **Conference**   City **São José do Rio Preto, SP**

Date **24/09/2018, 25/09/2018, 26/09/2018, 27/09/2018, 28/09/2018**
Modalities **Audio Description, Sign Language Interpreting, Subtitling for the Deaf and Hard of Hearing**

Accessometer

The Accessometer provides a world map of the legislation, standards and guidelines on media accessibility organised by countries.

Read more

LEGISLATION, STANDARDS AND GUIDELINES

Об утверждении Государственной программы о социальной защите и содействии занятости населения на 2016–2020 годы

Type **Legislation**   Year **2016**

Authoring **Government of Belarus**
Country **Belarus**

ГОСТ Р 57891-2017. НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ. ТИФЛОКОММЕНТИРОВАНИЕ И ТИФЛОКОММЕНТАРИЙ. Tiflokommentirovanie and tiflokommentariy. Terms and Definitions

Type **Standards**   Year **2017**

Authoring **Pavel Obiukh, Sergey Van'shin, Federal State Unitary Enterprise "Russian Scientific-**

TRANS MEDIA CATALONIA | UAB Universitat Autònoma de Barcelona

Disclaimer