Enhancing sensorial and linguistic accessibility with technology: further developments in the TECNACC and ALST projects

by Anna Matamala, Anna Fernández, and Carla Ortiz-Boix

5th International Conference Media for All. Audiovisual Translation: Expanding Borders. Dubrovnik, 25-27/09/2013.

Enhancing sensorial and linguistic accessibility with technology: further developments in the TECNACC and ALST projects

Presenters: Anna Matamala, Carla Ortiz-Boix

Co-author: Anna Fernández

Anna.matamala@uab.cat; Carla.Ortiz@uab.cat

CAIAC, Universitat Autònoma de Barcelona

With the financial support of the APOSTA project TECNACC (APOSTA2011-10) and the Spanish Ministerio de Ciencia e Innovación (project ref. code FFI2012-31024)

TECNACC & ALST projects

TECNACC	Post-edited MT vs human translation vs creation in AD	Evaluation of TTS AD's acceptance in films (Cat) as compared to human voices
ALST	Post-edited MT vs human translation in VO/OD	Evaluation of TTS VO/off-Screen dubbing in documentaries (Spanish)

TECNACC & ALST projects

- Common features
  - Oral transfer modes
  - Narrator with a clear voicing
    - off-screen dubbing
    - audio description
- Differences
  - MT engine
    - trained engine
    - online free engine
  - Language
    - Catalan
    - Spanish

TECNACC & ALST projects

- Previous or ongoing research/practice
  - TTS AD (Poland, Switzerland, Japan), TTS AST
  - MT in AVT: focus on subtitling in many projects such as SUMAT
- Novelty (to the best of our knowledge)
  - MT in oral modes such as AD and VO/Off-screen dubbing of documentaries
  - TTS acceptance in AD in Catalan
  - TTS in documentaries

TECNACC: TTS AD

TTS: 2-level experimental design

- Test on TTS AD acceptance
  - Female natural vs female artificial voice
  - Male natural vs male artificial voice

  By blind and visually impaired users

- Pre-test to select the voices to be used in the experiment

Voices

- Artificial voices: from the available engines in Catalan we selected 5 male and 5 female voices
  - Loquendo: Jordi and Montserrat
  - Verbio: Oriol and Meritxell
  - Acapela: Laia
  - iSpeech: Anna
  - FestCat: Pep, Jan, Teo and Ona

- Natural voices: 5 male and 5 female voices (with the help of ECAD)

AD units selection

- Viswanathan & Viswanathan (2005: 65) "**The same set of sentences** is used as input to all of the synthesizers in a study"
- 1 participant would therefore listen to the same set of AD units 20 times => fatigue and learning effect
- Two different sets of AD units: one set for the feminine voices and another set for the masculine ones => same stimulus for all feminine/masculine voice samples to be able to compare them properly
- Balance between the two sets of sentences as far as number of characters is concerned.

Voice selection test

- Pilot test
  - 6 participants
- Test
  - 20 participants
    - **6 men and 14 women ranging from 19 to 51**
- Two sessions:
  - 1st. synthetic voices
  - 2nd. natural voices (otherwise too long)
  - Short recorded AD units

Questionnaire's items selection

- Items related to ITU's quality factor rather than intelligibility – **9 items**
  - overall impression
  - listening effort / ease of listening / comfortable to listen to for a long period
  - acceptance
  - accentuation
  - pronunciation
  - speech pauses / audio flow / prosody
  - intonation
  - naturalness
  - voice pleasantness

Results

- Pilot test and experiment results match

1. Best feminine natural voice: D (**professional voice talent**)

2. Best masculine natural voice: no statistically significant differences among F, H and I, but H has a higher median and mean in overall impression (**student voice talent**)

3. Best feminine synthetic voice: no statistically significant differences between A and C (**Laia by Acapela and Meritxell by Verbio**), but A has a higher median and mean in overall impression (**Laia by Acapela**)

4. Best masculine synthetic voice: H (**Oriol by Verbio**)

The experiment

- Pilot test
  - 2 participants

- Test
  - 67 blind and visually impaired people
    - with many thanks to the **Visual Impairment Association of Catalonia**

The experiment

Age	Total	Women	Men
20 to 34	9	5	4
35 to 44	12	5	7
45 to 54	15	7	8
55 to 64	15	9	6
65 to 74	13	9	4
Older than 80	3	2	1
TOTAL	67	37	30

The experiment

- 2 male voices (synthetic/natural), 2 female voices (s/n)
- Same questionnaire and post-questionnaire
- Two different clips: female/male
- In depth analysis of the film, AD script and AD units to get to a balance between the two clips as far as:

The experiment

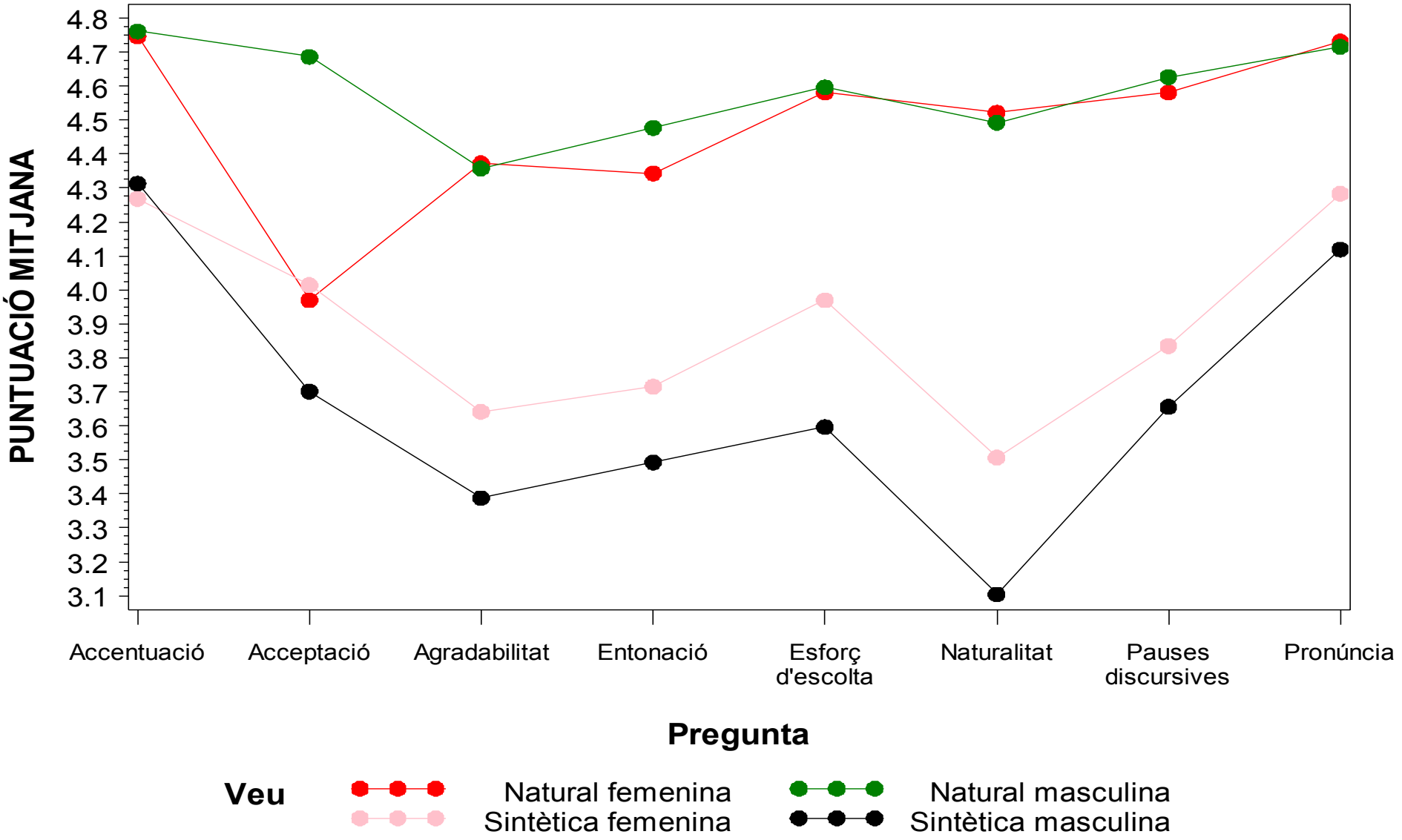
	Clip 1	Clip 2
Length in minutes	3 minutes	3 minutes and 6 seconds
Intervening characters	Anna and Dan	Anna and Dan
Background music	Opera *Così fan tutte ossia La scuola degli amanti*, by Mozart	Opera *Così fan tutte ossia La scuola degli amanti*, by Mozart
No potentially distracting and/or offensive content (eg. sexual tension)	Neutral	Neutral
AD density in characters	571 characters	537 characters

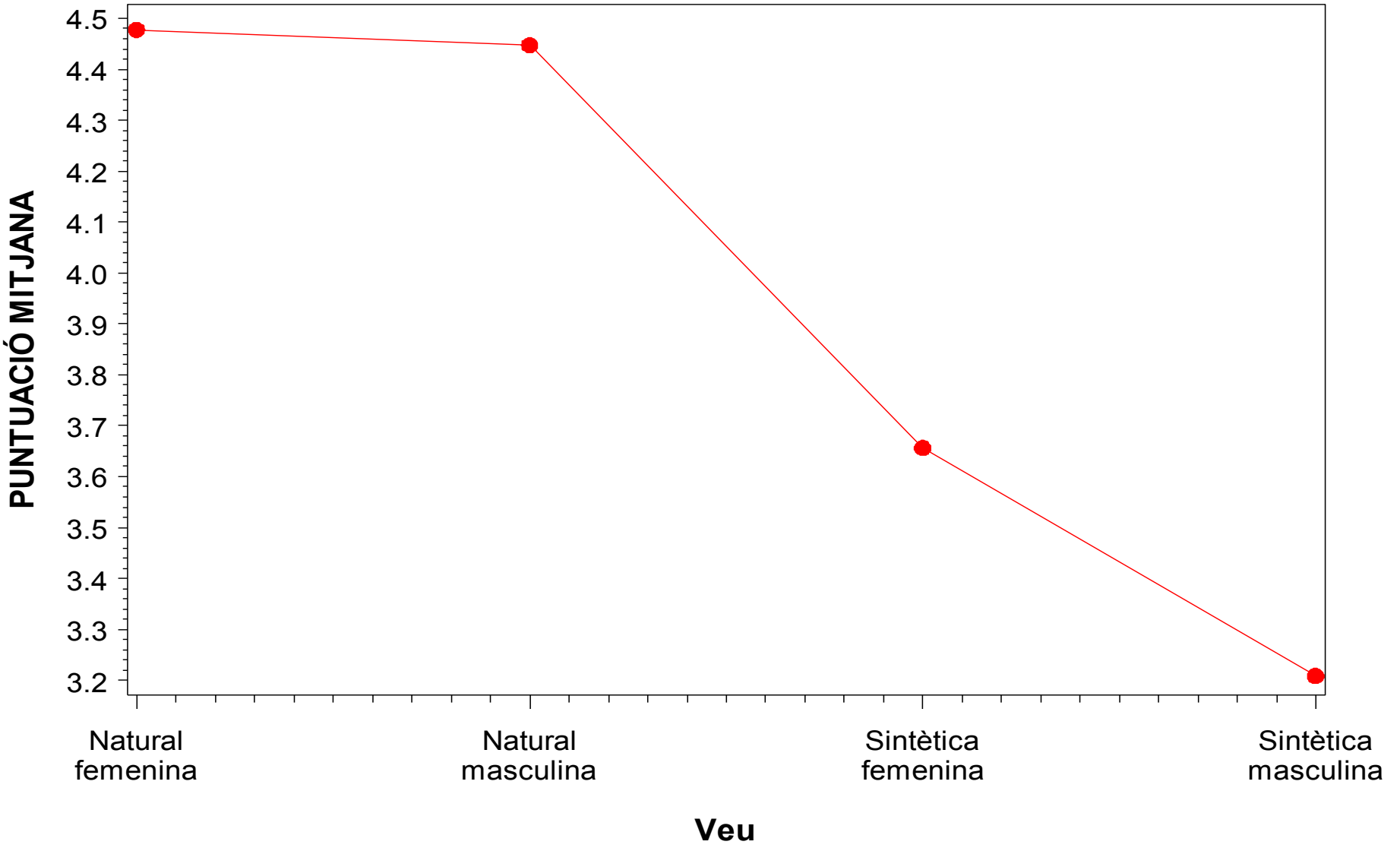
Results TTS AD

- Natural voices have higher values than artificial voices, as expected.

- No statistical differences between male/female natural voices.

- Female artificial voice, higher values than male artificial voice.


PUNTUACIÓ MITJANA
4.8
4.7
4.6
4.5
4.4
4.3
4.2
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
Accentuació
Acceptació
Agradabilitat
Entonació
Esforç d'escolta
Naturalitat
Pauses discursives
Pronúncia
Pregunta
Veu
Natural femenina
Sintètica femenina
Natural masculina
Sintètica masculina


PUNTUACIÓ MITJANA
4.5
4.4
4.3
4.2
4.1
4.0
3.9
3.8
3.7
3.6
3.5
3.4
3.3
3.2
Natural femenina
Natural masculina
Sintètica femenina
Sintètica masculina
Veu

Results TTS AD

- Overall impression/acceptance: mean higher than 3.2

- 94% participants (out of 67)
  - TTS is an alternative acceptable solution, although not the preferred one.

- Current work: MT in AD

ALST
Experiment 1: MT in the translation of documentaries

Carla Ortiz-Boix's PhD

- Documentaries:
  - Off-screen dubbing
  - Voice-over

- Machine translation vs human translation

- Also, if possible, TTS in Documentaries

First decisions

- What do I want to test?
  - MT + Post-editing vs human translation in terms of
    - Process
      - Time
      - Effort
        - » Technical (keyboard strokes)
        - » Cognitive (post-questionnaires)
    - Result
      - Quality assessment

First decisions

- Free online or trained engine
  - Customized engine: Domain specific
  - 90 documentaries (200,000 aprox words)
    - Challenging, time-consuming… exploring other possibilities.
- Participant profile:
  - professional translators vs MA students
    - influence of previous background
- Platform and integration of audiovisual content

Future work

- Define and carry out the post-editing experiments on Documentaries (2014)

- Carry out the TTS experiment on documentaries (2015)

Thank you

Enhancing sensorial and linguistic accessibility with technology: further developments in the TECNACC and ALST projects

Presenters: Anna Matamala, Carla Ortiz-Boix

Co-author: Anna Fernández

Anna.matamala@uab.cat; Carla.Ortiz@uab.cat

CAIAC, Universitat Autònoma de Barcelona

With the financial support of the APOSTA project TECNACC (APOSTA2011-10) and the Spanish Ministerio de Ciencia e Innovación (project ref. code FFI2012-31024)

Enhancing sensorial and linguistic accessibility with technology: further developments in the TECNACC and ALST projects

by Anna Matamala, Anna Fernández, and Carla Ortiz-Boix

5th International Conference Media for All. Audiovisual Translation: Expanding Borders. Dubrovnik, 25-27/09/2013.