

# Technology and AD: the TECNACC project



**Author and speaker: Anna Matamala**

**Co-authors: Anna Fernández-Torné,  
Carla Ortiz-Boix.**

**Languages and the Media conference, 9<sup>th</sup>  
International Conference on Language Transfer in  
Audiovisual Media. Berlin, 21-23/11/12.**

**Funded by APOSTA 2011-10 project.**

# Aim

- Present ongoing research related to the TECNACC project (funded by the APOSTA Prize to Young Researchers at UAB) in order to get feed-back and suggestions from colleagues

# Overview

- Overall aim and background:
  - why this project?
  - previous and ongoing research
- Exploratory research by Carla Ortiz-Boix: results and setbacks (MA, CA>ES)
- Ongoing research by Anna Fernández (PhD, EN>CA)
- Future prospects

# Overall aim

- Can existing language and translation technologies be successfully incorporated into AD creation?

# Why this project?

- Technology > increase accessibility?
  - Not human substitute but help
  - A way to increase the number of AD products
- Machine translation & speech synthesis
  - Explore a new application, not engine development
  - Expand on existing research

# TECNACC: MT & SS

- Machine Translation for AD (En>Ca)
  - Result & effort (translators' point of view)
- TTS AD in Catalan
  - Audience satisfaction (acceptability)
- Small-scale exploratory research on CA>ES (multilingual countries)

# Previous research

- Translating AD, is it possible?
  - Rosa Vallverdú, in Matamala & Orero (2007): not sure about this
  - Hycks: adaptation needed
  - Vercauteren & Remael (2010): it will increase because less cost

But no reference to the best of our knowledge to MT in AD.

# Previous research: MT

- In Audiovisual Translation, research on MT, has focussed on subtitling:
  - Popowich et al (2000)
  - Previous projects such as MUSA (2002-2004), E-title (2003-2005)
  - O'Hagan (2003), Armstrong et al. (2006)
  - Volk (2008), Volk & Harder (2007), Volk and Hardmeier (2009)
  - Specia, De Sousa (2011)
  - Ongoing projects SUMAT, SAVAS, etc.



# Previous research: MT

- Into Catalan/from Catalan
  - Many universities (for instance, UOC) and companies
  - Catalan newspapers: El Periódico, La Vanguardia, El Segre

# Previous research: TTS AD

- Polish TTS AD project:
  - Szarkowska (2011): Polish feature film
  - Walczak & Szarkowska for children
  - Szarkowska & Jankowska's analysis of "Volver" (AST).
  - Documentary: work in progress
  - Foreign feature film with Polish dubbing (MA, PhD)
- Conclusions in 3/5 stages: 94% viewers accept TTS AD as an interim solution and 63% are willing to accept it as a permanent solution.
  - Pro: quick access, low cost, no help from sighted friends needed.
  - Con: requires media literacy, improvement needed for TTS software and does not promote integration (although it is considered a complement, not a replacement)

# Previous research: TTS AD

- Spoken subtitles (Verboom et alii 2002)
- Swiss TXT (Juan Martínez, Gion Linder, Beatrice Caruso)
- UAS (Universal Accessibility System) at CAIAC (Orero & Serrano): TTS AD + subtitling + spoken subtitles

# Why can it be successful?

- Because AD is a text genre with specific features which can make it more prone to MT (short sentences, repeated structures) although some issues might be an obstacle (adjectivation, isochrony)
- Because previous successful experiences in MT, especially in CAT<>ES
- Because blind audiences are used to artificial voices
- Because we don't want to propose a fully automatic solution but want to explore the feasibility of semi-automatized processes and evaluate the professionals/ audience's response

# Exploratory research (Ortiz)

- MA thesis (3 months), CA>ES, MT + TTS AD
  - Catalan AD of
    - New series: “Gran Nord”. First chapter, 56 minutes.
    - Film: “Bruc”. 10 + 10 + 10 minutes (of 55 minutes).
  - Video file with audio + written scripts (special thanks to Rosa Vallverdú, TVC)

# Exploratory research (Ortiz)

- MT of AD from Catalan into Spanish
  - Engines
    - Apertium (open source)
    - Google Translate
  - Error categorization (human judge)
    - Based on Font Llitjós et alii (2005)
      - Missing word
      - Extra word
      - Wrong word order
      - Incorrect word
      - Wrong agreement

# Preliminary results on MT

## “Gran Nord”

### GOOGLE TRANSLATE/APERTIUM

- 1687 words in Catalan > 1752 / 1808 in Spanish (157 sentences).
- 62 / 160 mistakes (3.52% / 8.85%)
  - Missing word > 2 / 1
  - Extra word > 0 / 0
  - Wrong word order > 16 / 61
  - Incorrect word
    - » Language errors > 22 / 33
    - » Contextual errors > 9 / 17
  - Wrong agreement > 13 / 38
- 109 / 57 sentences without mistakes (69.42% / 36.31%) > 48 / 100 sentences with mistakes (30.57% / 63.69%)

# Preliminary results on MT

“Bruc”

## GOOGLE TRANSLATE/APERTIUM

- 2697 words in Catalan > 2804 / 2791 in Spanish (285 sentences).
- 192 / 346 mistakes (6,85% / 12,40%)
  - Missing word > 0 / 4
  - Extra word > 0 / 0
  - Wrong word order > 113 / 114
  - Incorrect word
    - » Language errors > 20 / 78
    - » Contextual errors > 13 / 30
  - Wrong agreement > 37 / 85
- 162 / 108 sentences without mistakes (56,84% / 37,89%) > 123 / 177 sentences with mistakes (43,157% / 62,11%)



# Evaluating TTS AD

- Selecting and testing the artificial voice
  - Verbio ([Carlos](#)), Festival (Uvigo), OS X ([Diego](#)), Acapela ([Antonio](#)), OS X ([Monica](#)), Acapela ([Inés](#)), Natural voices (no/low/high distortion)
- Opinion measures, rather than TTS performance (intelligibility tests)
- Items to be scored according to ITU (1994)



<b>Overall impression</b>  How do you rate the quality of the sound of what you just heard?  <input type="radio"/> Excellent <input type="radio"/> Good <input type="radio"/> Fair <input type="radio"/> Poor <input type="radio"/> Bad	<b>Listening effort</b>  How would you describe the effort you were required to make in order to understand the message?  <input type="radio"/> Complete relaxation possible; no effort required <input type="radio"/> Attention necessary; no appreciable effort required <input type="radio"/> Moderate effort required <input type="radio"/> Considerable effort required <input type="radio"/> No meaning understood with any feasible effort	<b>Pronunciation</b>  Did you notice any anomalies in pronunciation?  <input type="radio"/> No <input type="radio"/> Yes, but not annoying <input type="radio"/> Yes, slightly annoying <input type="radio"/> Yes, annoying <input type="radio"/> Yes, very annoying
---	---	--

<b>Speaking rate</b>  The average speed of delivery was:  Just right <input checked="" type="radio"/>  Slightly slow <input type="radio"/> Slightly fast  Fairly slow <input type="radio"/> Fairly Fast  Very slow <input type="radio"/> Very fast  Extremely slow <input type="radio"/> Extremely fast	<b>Voice pleasantness</b>  How would you describe the voice?  <input type="radio"/> Very pleasant <input type="radio"/> Pleasant <input type="radio"/> Neutral <input type="radio"/> Unpleasant <input type="radio"/> Very unpleasant	<b>Voice naturalness</b>  Did the voice sound natural?  <input type="radio"/> Very natural <input type="radio"/> Natural <input type="radio"/> Neutral <input type="radio"/> Unnatural <input type="radio"/> Very unnatural
---	---	---

<b>Ease of listening</b>  Would it be easy to listen to this voice for long periods of time?  <input type="radio"/> Very easy <input type="radio"/> Easy <input type="radio"/> Neutral <input type="radio"/> Difficult <input type="radio"/> Very difficult	<b>Comprehension problems</b>  Did you find certain words hard to understand?  <input type="radio"/> Never <input type="radio"/> Rarely <input type="radio"/> Occasionally <input type="radio"/> Often <input type="radio"/> All of the time	<b>Articulation</b>  Were the sounds distinguishable?  <input type="radio"/> Yes, very clear <input type="radio"/> Yes, clear enough <input type="radio"/> Fairly clear <input type="radio"/> No, not very clear <input type="radio"/> No, not at all	<b>Acceptance</b>  Do you think that this voice could be used for an interactive telephone system or a handheld device?  <input type="radio"/> Yes <input type="radio"/> No
---	--	---	--



# Preliminary results

- Best rated voices (accepted for AD):
  - Natural voice (no distortion)
  - Acapela (male): 3,78
  - Verbio (male) and Acapela (female): 3,39
- Initial research with some methodological problems (testing ground)
- Future research in MT/TTS VO

# PhD thesis (Fernández)

- From English into Catalan,
- TTS AD (April 2012-March 2013), MT AD (April 2013-April 2014).
- Test the reception of TTS AD in comparison with human-voiced AD

# Steps

- Test 1: voice selection (natural/artificial)
  - Selecting the voices
  - Selecting the evaluation system
  - Carrying out a pilot test (6 participants)
  - Voice selection experiment (20 participants)
- Test 2: TTS AD application (sample to be defined- with help from Statistical Service at UAB)

# Selecting the voices

- 5 MALE/5 FEMALE x artificial/natural = 20 voices
- Artificial voices> available engines in Catalan: Loquendo (Jordi/Montserrat), Verbio (Oriol/Meritxell), Acapela (Laia), Nuance (Núria), VozMe (fem/male), Ispeech (fem), eSpeak (male/female). Initial selection to 5 female + 5 male (ongoing).
- Natural voices> professional voice talents (special thanks to Iola Ledesma, from the Escola Catalana de Doblatge). 5 female + 5 male (recordings on Friday)

# Selecting the excerpts

- Film (miscellaneous): “Closer” (2004)
- Already audiodescribed in Catalan (250 units).
- Selection of AD units for test 1 (voice selection): random selection based on number of characters
- Selection of AD segment for test 2 (not done)

# Selecting the evaluation items

- ITU (1994):
  - overall impression, listening effort, pronunciation, speaking rate, voice pleasantness, comprehension problems, articulation, acceptance.
- Viswanathan & Viswanathan (2005), for interactive telephone or wireless handheld information :
  - overall impression, listening effort, pronunciation, speaking rate, pleasantness, naturalness, audio flow, smooth, ease of listening, comprehension problems, articulation, acceptance



# Selecting the evaluation items

- Hinterleitner et al (2011), for audiobooks
  - Overall impression, voice pleasantness, speech pauses, accentuation, intonation, emotion, listening effort, comprehension problems, acceptance.
- Cryer, Home and Wilkins for RNIB (2010), for AD, not validated yet
  - rating the voice and application in terms of overall impression, pleasantness, comprehension, pronunciation, prosody, comfortable to listen to for a long period, responsiveness, speaking rate, naturalness, listening effort, appropriate tone

# Selecting the evaluation items

- Limit number of questions
- Compare formulation in different [questionnaires](#)
- Formulate the questions in Catalan:
  - overall impression, accentuation, pronunciation, pleasantness, naturalness, intonation, speech pauses, listening effort, acceptance.

# Selecting the evaluation items

- Post-test questions based on Walczak (2010), Maczynska (2011), Chmiel & Mazur (2012), and Pazos (2012)
  - AD usage (how often? What?), TTS usage, preference for male/female voices, TTS AD as an additional solution (for specific products?)

# Methodological considerations

- December: pilot test (voice selection) and sampling Catalan blind population
- January/February: experiment + analysis

# Future prospects

- Finish TTS AD experiment
- Initiate MT AD experiment (defining evaluation system). Any suggestions would be welcome!
- Process integration in a platform
- Expanding this ecosystem to other transfer modes, namely voice-over

THANK YOU!

anna.matamala@uab.cat