

Technology for accessibility in multilingual settings: the way forward in AD?

Anna Fernández, Anna Matamala, Carla Ortiz

**The translation and reception of multilingual films.
Univ. Montpellier 3, 15-16/06/2012.**

Funded by APOSTA 2011-10.



Technology for accessibility in multilingual settings: the way forward in AD?

Anna Fernández/Anna Matamala/Carla Ortiz
anna.matamala@uab.cat

Universitat Autònoma de Barcelona

CAIAC

Funded by the APOSTA Prize- UAB

Overview

- The TECNACC project (Aposta Prize for young researchers at UAB)
- Previous research
- The project: ongoing research
- Preliminary results
- Future prospects

TECNACC

- Technology > increase accessibility?
 - Not to substitute human translators but to help them maybe?
 - To offer AD in products which would not have AD?
- Machine translation & speech synthesis
 - Explore a new application, not develop MT engines or tools

TECNACC: MT & SS

- Machine Translation for AD (En>Cat, Cat>Spanish in multilingual countries)
 - Result & effort (translators' point of view)
- TTS AD in Catalan/ in Spanish
 - Audience satisfaction (acceptability)
 - Audience comprehension

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 focused on subtitling:
 - Popowich et al (2000)
 - MUSA (2002-2004)
 - O'Hagan (2003)
 - E-title (2003-2005)
 - Armstrong et al. (2006)
 - Volk (2008), Volk & Harder (2007), Volk and Hardmeier (2009)
 - Ongoing project SUMAT (2011-2014)
 - Specia, De Sousa (2011)

Previous research: MT

- Into Catalan/from Catalan
 - Universities such as UOC
 - Catalan newspapers: El Periódico, La Vanguardia

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

- Universal Accessibility System (UAS) at CAIAC (Orero, Serrano et alii)
 - 2011/12: 5 films
 1. Subtitling: Multilanguage subtitling and SDH
 2. Audio description: Multilanguage AD
 3. Spoken subtitles: through speech synthesis: subtitles --> voiced subtitles
 4. Automatic AD: through speech synthesis: AD text --> voice to create AD in an automatic mode
 5. Delivery of spoken subtitles and the whole performance or AD through VoIP for those who use hearing aids via Wifi.
 6. Emergency pack - which adds a pre-recorded sign language for some emergency messages to all these previous services. The emergency will also activate the vibration mode on the mobile phone, to alert deaf users to the incoming information.

Previous research: TTS AD

- Spoken subtitles in the Netherlands (Verboom et alii 2002)
- Studies on the application of speech technologies for the blind (Freitas & Kouropetroglou 2008, Papadopulos 2009)
- Swiss TXT (Juan Martínez, Gion Linder)

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

Ongoing research

- What MT system to choose? Available systems CAT and (ES or EN)
 - Apertium <http://www.apertium.org/?id=translatetext>
 - Google Translate <http://translate.google.cat/>
 - Lucy Kwick Translator
<http://www.lucysoftware.com/catala/traducci-automtica/lucy-It-kwik-translator-/lucy-It-quick-translator.html?parent=&subid=>
 - OpenTrad <http://www.opentrad.com//index.php?idioma=ca>
 - Instituto Cervantes
http://www.cervantes.es/lengua_y_ensenanza/tecnologia_espanol/informacion.htm
 - InterNostrum <http://www.internostrum.com/>
 - ITranslate4EU <http://itranslate4.eu/>
 - LexPress <http://www.standing.com/trad/>
 - Softcatalà <http://www.softcatala.org/traductor>
 - Salt (Valencian dialect) http://www.cefe.gva.es/polin/val/salt/apolin_salt.htm
 - N- II (UPC): <http://www.n-ii.org/>

Ongoing research

- How to evaluate machine translation output?
 - Automatic metrics: BLEU, NIST, TER, METEOR, GTM
 - Confidence scores in MT systems (Blatz et al. 2004, Specia et alii 2009, Bach et alii 2011)
 - Manual evaluation (human judges)
 - Intelligibility/fluency
 - Fidelity/adequacy (meaning)
 - Post-editing effort (O'Brien 2011, Krings 2001)
 - Task time
 - Cognitive effort (eye-tracking/TAP)
 - Technical effort (keyboarding)

Ongoing research

- How to evaluate TTS AD?
 - Not interested in an evaluation of TTS performance (intelligibility tests using objective/acoustic metrics based on statistic methods) but an evaluation of a specific application by means of...
 - Performance measures: comprehension of TTS AD
 - Opinion measures: user's satisfaction/ quality of experience > “acceptability”

Ongoing research

- How to evaluate comprehension of TTS AD?
 - Experiment based on Cabeza-Cáceres Phd thesis
- How to evaluate acceptability?
 - Mean Opinion Score (MOS), 10 to 30 listeners, scale of 1 to 5 (1=Bad, 2=Poor, 3= Fair, 4=Good, 5= Excellent).
 - Items to be scored according to ITU (1994): global impression, listening effort, comprehension problems, speech sound articulation, pronunciation, speaking rate, voice pleasantness. Adaptation of Ojala (2006) to our project?



<p>Overall impression</p> <p>How do you rate the quality of the sound of what you just heard?</p> <ul style="list-style-type: none"> <input type="radio"/> Excellent <input type="radio"/> Good <input type="radio"/> Fair <input type="radio"/> Poor <input type="radio"/> Bad 	<p>Listening effort</p> <p>How would you describe the effort you were required to make in order to understand the message?</p> <ul style="list-style-type: none"> <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 	<p>Pronunciation</p> <p>Did you notice any anomalies in pronunciation?</p> <ul style="list-style-type: none"> <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
---	---	--

<p>Speaking rate</p> <p>The average speed of delivery was:</p> <p>Just right <input checked="" type="radio"/></p> <p>Slightly slow <input type="radio"/> Slightly fast <input type="radio"/></p> <p>Fairly slow <input type="radio"/> Fairly Fast <input type="radio"/></p> <p>Very slow <input type="radio"/> Very fast <input type="radio"/></p> <p>Extremely slow <input type="radio"/> Extremely fast <input type="radio"/></p>	<p>Voice pleasantness</p> <p>How would you describe the voice?</p> <ul style="list-style-type: none"> <input type="radio"/> Very pleasant <input type="radio"/> Pleasant <input type="radio"/> Neutral <input type="radio"/> Unpleasant <input type="radio"/> Very unpleasant 	<p>Voice naturalness</p> <p>Did the voice sound natural?</p> <ul style="list-style-type: none"> <input type="radio"/> Very natural <input type="radio"/> Natural <input type="radio"/> Neutral <input type="radio"/> Unnatural <input type="radio"/> Very unnatural
--	---	---

<p>Ease of listening</p> <p>Would it be easy to listen to this voice for long periods of time?</p> <ul style="list-style-type: none"> <input type="radio"/> Very easy <input type="radio"/> Easy <input type="radio"/> Neutral <input type="radio"/> Difficult <input type="radio"/> Very difficult 	<p>Comprehension problems</p> <p>Did you find certain words hard to understand?</p> <ul style="list-style-type: none"> <input type="radio"/> Never <input type="radio"/> Rarely <input type="radio"/> Occasionally <input type="radio"/> Often <input type="radio"/> All of the time 	<p>Articulation</p> <p>Were the sounds distinguishable?</p> <ul style="list-style-type: none"> <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 	<p>Acceptance</p> <p>Do you think that this voice could be used for an interactive telephone system or a handheld device?</p> <ul style="list-style-type: none"> <input type="radio"/> Yes <input type="radio"/> No
---	--	---	--



Ongoing research

- What TTS system to choose? Some available systems (CAT & ES)
 - Festival <http://www.talp.upc.edu/festcat/>
 - Acapela <http://www.acapela-group.com/text-to-speech-interactive-demo.html>
 - Verbio <http://www.verbio.com/webverbio3/index.php/es/demo-separator/demo-tts-online.html>
 - VozME <http://vozme.com/index.php?lang=ca>
 - eSpeak <http://espeak.sourceforge.net/samples.html>
 - TALP-OGMIOS <http://www.talp.cat/ttsdemo/index.php>
 - Loquendo/Nuance <http://www.loquendo.com/en/demo-center/interactive-tts-demo/>
 - CereProc http://www.cereproc.com/support/live_demo

Preliminary results

- Carla Ortiz-Boix's research: explore feasibility with a small corpus
 - Catalan AD of
 - New series: “Gran Nord”. First chapter, 8:38 minutes.
 - Film: “Bruc”. 10 + 10 + 10 minutes (of 55 minutes).
 - Video file with audio
 - Written scripts (thanks to R. Vallverdú, TVC)

Preliminary results

- MT of AD from Catalan into Spanish
 - Engines
 - Apertium
 - Google Translate
 - Error categorization
 - 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%)

Preliminary results

- TTS AD: exploratory reception study (ongoing)
 - Example [Verbio](#)
 - Example [Festival](#)
 - Example [Acapela Inés](#)
 - Example [Acapela Antonio](#)

Future prospects

- Continue with the experiments with a wider corpus/
number of participants/other language combinations
- MT AD:
 - Possible triangulation of data from keylogging & eye-tracking to
assess post-editing effort?
 - Comparing effort and results in
 - AD creation
 - AD human translation
 - AD MT + human post-editing
- TTS AD:
 - Compare reception of speech synthesis/standard AD/ dubbed
AD
- Process integration in a platform

Technology for accessibility in multilingual settings: the way forward in AD?

Anna Fernández/Anna Matamala/Carla Ortiz
anna.matamala@uab.cat

Universitat Autònoma de Barcelona

CAIAC

Funded by the APOSTA Prize- UAB