Live subtitling with speech recognition: causes and consequences of text reduction

The textual complexity of audio-description with audio-subtitling

Making multilingual films accessible for the blind and partially sighted

Aline Remael

Université Antwerpen

Languages and the Media, Berlin, 21-23 November 2012

• Multilingualism in film
• The AD & AST combination
• The corpus of three films (clips)
• Project research questions/hypotheses
• Analysis of three scenes
• Concluding thoughts & questions

Multilingualism in film

• Today's society is multilingual.
• This is reflected in today's cultural products, including film.
• Films opt for multilingualism for various reasons:
  • realism
  • narratological/ideological issues
  • financial backing and/or international appeal
(see also Díaz Cintas 2011)

The AD & AST combination (1)

Accessibility:
• Multilingual films are a challenge
  • Audio-description (AD) & Audio-subtitling (AST)
(see also Braun & Orero, 2011; Remael, 2012)

The corpus of three films

• The Ban Advocates. From Victims to Champions
  (Anderson 2009)
• Tirza (van den Berg 2010)
• Süskind (van den Berg, 2011)

Accessible productions: collaboration Suzanne Verberk (www.renero.be), Mervin van der Heijden (www.soundfocus.nl), partners in the present research project (see also Remael, 2012).

Project research questions/hypotheses

1. Film dialogues or interviews function in a narratological context. They determine and are determined by the filmic narrative (Remael 2012).
  OR
  • Narrative movement and story are (co-)determined by dialogic interaction of different speakers and their (fictional) agenda.
  • Complexity of this dialogic interaction has an influence on strategy that works best to achieve accessibility.

Aline Remael (@artesis.be)
Bieke Luyckx, Luuk Van Waes & Mariëlle Leijten (@ua.ac.be)
Live subtitling with speech recognition: causes and consequences of text reduction

Project research questions/hypotheses

2.
- Multilingualism has a specific narrative function.
- Modifying multilingualism modifies dialogic interaction and the film story.
- AD and AST decreases and adds to the complexity of a scene/film.

Project research questions/hypotheses

3.
- Making films accessible is an added cost.
- Synthetic voices for AD and AST may be only solution.
- The “voice” of the industry makes itself heard in the film.
- Relate decisions for/against synthetic voices to variables?

Project research questions/hypotheses

4. Further research
- Film dialogues work on two levels (Remael 2003).
- How do the translational shifts affect communication?
- Does the target audience feel these shifts matter?
- Which solutions work best?

Three clips

I. Complexity of the dialogic interaction
The Ban Advocates: interviews
- Textual complexity of dialogic interaction limited:
  → narration
  → interviews
- Interaction is genre-bound
- However, interviewees have a voice of their own: English, Dari, Vietnamese, Tigrinya and Serbian, ....

2. AD & AST impact on (voices in) the story
The Ban Advocates
- Voices read by one male AD/AST & one female AST
- AD identifies the speakers.
- Subtitles are read as written
- AST resembles the voice of “authority”
- AST is narration, similar to the AD
- However, even original interviews are structured speech
3. Synthetic voices - filmic variables

The Ban advocates:

• **Content** outweighs formulation & intonation
• **Interaction** is limited and regulated
• The ‘natural’ AST voices sound *much like* AD narration
• Synthetic (quality) voices are *unproblematic*

1. The complexity of the dialogic interaction

Tirza: acting and voice-over

• **Complex** emotional scene even though only one speaker
• Conveys a *psychological state* as much as more than content
• Irregular speech rhythms, hesitations and code-switching

2. AD & AST impact on (voices in) the story

Tirza:

• AD the usual narrative tone ↔ AST
• Original dialogue, heard in between the ‘acted’ AST
• Both retain rhythm, emotionality and code switching
• Character does not lose his voice

**HOWEVER:**

• Complexity is added to complexity
• Does this remain comprehensible?

3. Synthetic voices - filmic variables

Tirza

• How Jörgen speaks is as important as what he says
• Dialogue and AST mix well
• In other scenes interaction is added
• **Contrast** narrative AD ↔ AST/dialogue mix is marked

**ALTERNATIVELY**

• What would the mix of dialogue and more neutral synthetic AST yield?
• AD and AST still distinguishable? Can technology handle this?
• Synthetic simplification preferable to natural complexity?

Same three questions – different solution (1)

Süskind: acting and dubbing

• Film: German and Dutch/German
• Accessible film: All perfect Dutch
  → Unusual for a “subtitling country”
  → Simplification that may add complexity
• Rewritten Dutch AST functions like film dialogue
• **Difference** AST narrative AD obvious.
• **No complexity** of “added” voices, multilingualism lost

Same three questions – different solution (2)

Issues

• Ideological, realistic and narrative function of German lost
• **Character identification** through acting/intonation only
• Added cost of rewriting subtitles
• Are synthetic voices better at “clean” voices?
Live subtitling with speech recognition:
causes and consequences of text reduction

**Concluding thoughts**

- Voice of AD = a “narrative voice” in all films
  - synthetic voices OK

- Interviews: The Ban Advocates = little emotionality, constant rhythms, hardly any interaction
  - synthetic voices OK

- Analysis on the basis of our hypotheses/research questions
  - identification of scenes and issues for testing

**Concluding questions: summary (1)**

**Starting point:** the ideological, narrative and realistic functioning of multilingualism in film is important

- Which type of recording is preferable for AD & AST: voice-over/dubbing, for audience/for industry?

- Can synthetic voices manage acting emotions, differing rhythms, intonation, interpersonal interaction?

- Is the complexity of “acted” AST added to/mixed with film dialogue complexity a good thing?

**Concluding questions: summary (2)**

- Do synthetic and neutral “narrative” AST voices work in complex scenes?

- Is such simplification better for comprehension?

- Do synthetic voices function better in dubbed alternative?

- Combination text to speech for AD + actors for AST?

**Further research**

Next stage = testing with target audiences

**HOWEVER**

- Will audience preferences be affordable?

- Would developers of Dutch text to speech software care to join the project?

**References**


**THANK YOU FOR YOUR ATTENTION**

QUESTIONS?

Aline Remael (@artesis.be)
Bieke Luyckx, Luuk Van Waes & Mariëlle Leijten (@ua.ac.be)