Quality in Text to Speech

“Text without meaning is but the ruin of the speech”

Antoine Kauffeisen
VP Marketing
Acapela Group
Quality in Text to Speech

Methodologies to test quality in text to speech
Quality in Text to Speech

Methodologies to test quality in text to speech

- Mean Opinion Score
- Pair comparison
- Accuracy testing
- Modified Rhyme Test
- Categorical estimation
- Phonetically balanced word lists
- Semantically Unpredictable Sentences
- Pairs Intelligibility Testing
- Prosody Unveiling Through Restricted Representation
- And some more
Quality in Text to Speech

Methodologies to test quality in text to speech

“The doors are closing, please stand clear of the doors.”

Voice A
Voice B
Voice C
Voice D
Voice E
Quality in Text to Speech

Methodologies to test quality in text to speech

“Damn, those alien bastards are gonna pay for shooting up my ride.”

Voice A
Voice B
Voice C
Voice D
Voice E
You can measure performance of the engine but does it say something about quality?
Does it say something about quality if you don’t know your application?
can 24,438 applications
have the same voice

and still talk quality?
Let's take another angle

Project Director, Dr. John Drew

www.buckingham.ac.uk/djo
www.dickens2012.org
What is the meaning of speech in your application?
find your way
What is the text made of?

- How can you best **describe the content** that will be transformed into speech?
- Dive into the meaning of your content
  - linguistic statistics,
  - messages,
  - specific words,
  - expressions,
  - styles, structures,
  - formats,
  - acronyms
  - etc....
Platform 4, for the 14:25 First Capital Connect service to Haddiscoe. A shuttle service is in operation between Merthyr Tydfil, and Llwynypia. Customers for Hackney Wick, should change at Gilschochill.

Please take care when boarding at Brondesbury Park, due to lightning having damaged equipment.

The train now standing at Platform 19A, is the 12:29, South Eastern service to Aberystwyth.
Tuning Rachel

- Voice already used in the application vs. custom voice
- Corpus analysis
- Additional application oriented recording added
- Specific application lexicon
- Specific equalization
What do you want the voice to convey?

- What **voice dependency** does your application have?
- **Did content analysis** raise a specific voice expectation?
- **Voice already existing?**
- Do you want your **users to choose the voice**?
- How would you **describe the voice you need**?
  - **Specific accents**
  - **Dialects**
  - **Characters**
  - **Celebrities**
  - **Emotions**
  - **Music**
  - **Audio chart**
  - **Sound design**
create your voice
your voice
following your specification for your application
your voice
tuned to your application

- Personalized script,
- Voice casting
- Voice direction
- Special effects
- Emotions & styles
- Sound Design
- and some more....

to get the best possible quality
Quality in Text to Speech

“thank you for your attention”

Antoine Kauffeisen
VP Marketing & Tiny Cars
Acapela Group
antoine.kauffeisen@acapela-group.com