

ATA 60TH ANNUAL CONFERENCE
OCT 23-26 2019 *Palm Springs* CALIFORNIA

Interlingual Real-Time Closed Captions: Where Accessibility Meets Translation



Pablo Romero-Fresco
p.romero-fresco@roehampton.ac.uk
UniversidadeVigo

Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- Realtime voice writing (intralingual respeaking)
 - Quality in realtime captioning
- Interlingual realtime closed captioning (interlingual respeaking)
 - Training, research and practice
 - Future developments

Where does realtime closed captioning sit?

Audiovisual Translation

Dubbing
1980s

Subtitling
1970s

Voice-over
1900s

Localisation
1980s

Where does realtime closed captioning sit?

Audiovisual Translation

Dubbing
1980s

Subtitling
1970s

Voice-over
1900s

Localisation
1980s

Media accessibility

Captioning



Subtitling for the deaf
and hard of hearing
1980s



Audio description
2000s



Video description

Where does realtime closed captioning sit?

Audiovisual Translation

Dubbing
1980s

Subtitling
1970s

Voice-over
1900s

Localisation
1980s

Media accessibility

Captioning

Subtitling for the deaf
and hard of hearing
1980s

Audio description
2000s

Video description

Realtime voice writing

Intralingual respeaking
2000s

Where does realtime closed captioning sit?

Audiovisual Translation

Dubbing
1980s

Subtitling
1970s

Voice-over
1900s

Localisation
1980s

Media accessibility

Captioning

Subtitling for the deaf
and hard of hearing
1980s

Audio description
2000s

Video description

Realtime voice writing

Intralingual respeaking
2000s

DRAFT INTERNATIONAL STANDARD
ISO/IEC DIS 20071-23

ISO/IEC JTC 1/SC 35

Secretariat: **AFNOR**

Voting begins on:
2017-04-28

Voting terminates on:
2017-07-20

**Information technology — User interface component
accessibility —**

Part 23:
**Guidance on the visual presentation of audio information
(including captions and subtitles)**

Technologies de l'information — Accessibilité du composant interface utilisateur —

Partie 23: Titre manqué

ICS: 35.240.20



-For those who **cannot hear or understand** the audio content

-Persons

with hearing loss

with learning difficulties or cognitive disabilities

watching a movie in a non-native language

who cannot hear the audio content due to environmental conditions

Media Accessibility

Access to media products, services, and environments for all persons who cannot, or cannot completely, access them in their original form (Greco, 2016b)

Realtime voice writing (intralingual respeaking)
as an example of the wide notion of media accessibility

Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

Comparison of live captioning methods

(Lambourne, 2007)

	Velo- type	Dual Keyboard	Steno	Voice writing
Delay	Mid	Mid	Low	Low
Speed	Mid	Mid/High	Very high	Mid/High
Accuracy	Mid	Mid/High	High	High
Difficulty	Mid	Mid	High	Mid
Cost	Mid	Mid	Mid/High	Low
Errors	Mid	Low	Mid	Mid

Comparison of live captioning methods

(Lambourne, 2007)

	Velo- type	Dual Keyboard	Steno	Voice writing
Delay	Mid	Mid	Low	Low
Speed	Mid	Mid/High	Very high	Mid/High
Accuracy	Mid	Mid/High	High	High
Difficulty	Mid	Mid	High	Mid
Cost	Mid	Mid	Mid/High	Low
Errors	Mid	Low	Mid	Mid

Current state of affairs

- Respeaking has overtaken stenography as the preferred form of live subtitling.
- The US and Canada mainly use stenography, but things are changing.
- Australia uses a mix, although respeaking is the most common method now.
- In Europe....



Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- Realtime voice writing (intralingual respeaking)

What is realtime voice writing?

“a technique in which a respeaker listens to the original sound of a (live) programme or event and respeaks it, including punctuation marks and some specific features for the deaf and hard-of-hearing audience, to a speech-recognition software, which turns the recognized utterances into subtitles displayed on the screen with the shortest possible delay”(Romero-Fresco 2011: 1)

Translation
Practices
Explained

SUBTITLING THROUGH SPEECH RECOGNITION: RESPEAKING

Pablo Romero-Fresco



Realtime voice writing as a profession

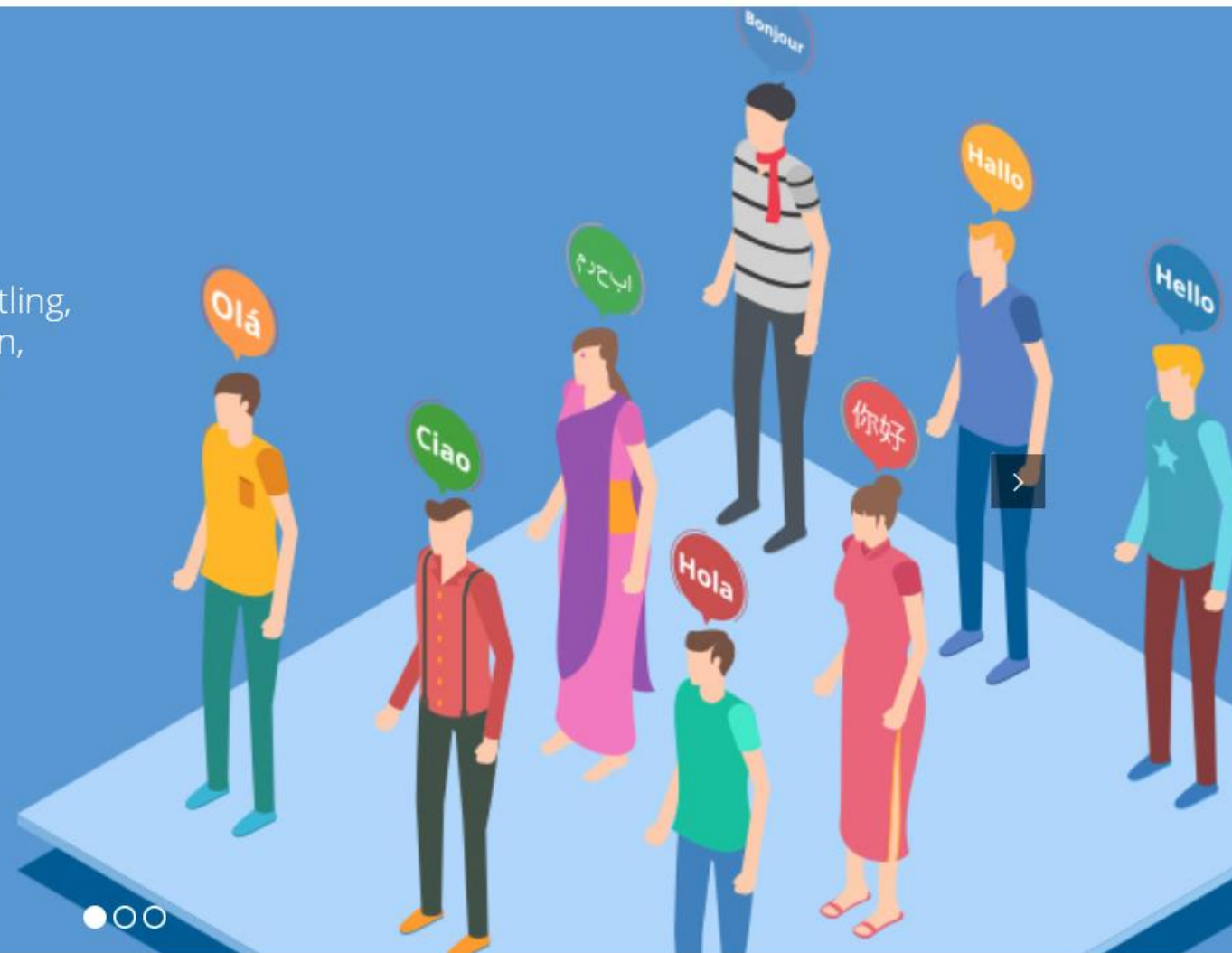
ILSA PROJECT

Interlingual Live Subtitling for Access

Helping to develop Interlingual Live Subtitling, a new discipline that combines translation, subtitling and simultaneous interpreting.



[Learn More](#)



INTERLINGUAL LIVE SUBTITLING FOR ACCESS (ILSA) IS A THREE-YEAR (SEPTEMBER 2017-AUGUST 2020) PROJECT CO-FINANCED BY THE EUROPEAN UNION UNDER THE ERASMUS+ PROGRAMME, KEY ACTION 2 - COOPERATION FOR INNOVATION AND THE EXCHANGE OF GOOD PRACTICES, KA203 - STRATEGIC PARTNERSHIPS FOR HIGHER EDUCATION (REFERENCE NUMBER 2017-1-ES01-KA203-037948). AS A STRATEGIC PARTNERSHIP PROMOTING OPEN AND INNOVATIVE

Realtime voice writing as a profession

- Highly educated women under 40 with a BA and language-related MA.
- Working as voice writers for 6 years and typically work for captioning companies (in-house or freelance).
- They do not normally respeak for more than one hour.
- Differences in equipment depending on country.
- Accuracy and delay is approached differently depending on country.

Training in realtime voice writing

- Training in realtime voice writing at university level is limited.
- Modules form part of larger MA programmes, such as AVT or Interpreting.

INTRODUCTORY TRAINING

- University of Leeds, UK
- Universitat Autònoma de Barcelona, Spain
- University of Parma, Italy

THOROUGH TRAINING

- University of Antwerp (6 months)
- University of Roehampton (3 months)

ONLINE TRAINING

- University of Vigo (CPD) (5 months)

Required skills

Description: Voice writing is to captioning what interpreting is to translation (a leap from the written to the oral without the safety net provided by time).

Process: interpreting

Result: captioning

Technology: speech recognition

Required skills

Interpreting

- Research
- Split attention
- Décalage
- Pressure management

Captioning

- Reformulation
- Spelling, punctuation
- Captioning issues: speaker identification, subtitle position...

Voice writing

- Software
- Fast and accurate dictation
- Multitasking

Knowledge of speech recognition

How to dictate: idea units

“See” = C or sea or see

“You have to see, for example, the advantages...”

=

“You have to C / sea / see, for example, the advantages...”

“You have to, for example, see the advantages...”

Multitasking

- Listening (ST)
- (Re)Speaking (TT)
- Listening (TT)
- Watching (ST and TT)
- Writing (corrected TT)

Multitasking

0 1 2 3 4 5 6

[watch and listen]

In 1998, the financial recession brought about the loss of hundreds of..

[watch, listen and respk]

In **1998**, the financial (recession has brought about the loss ...)

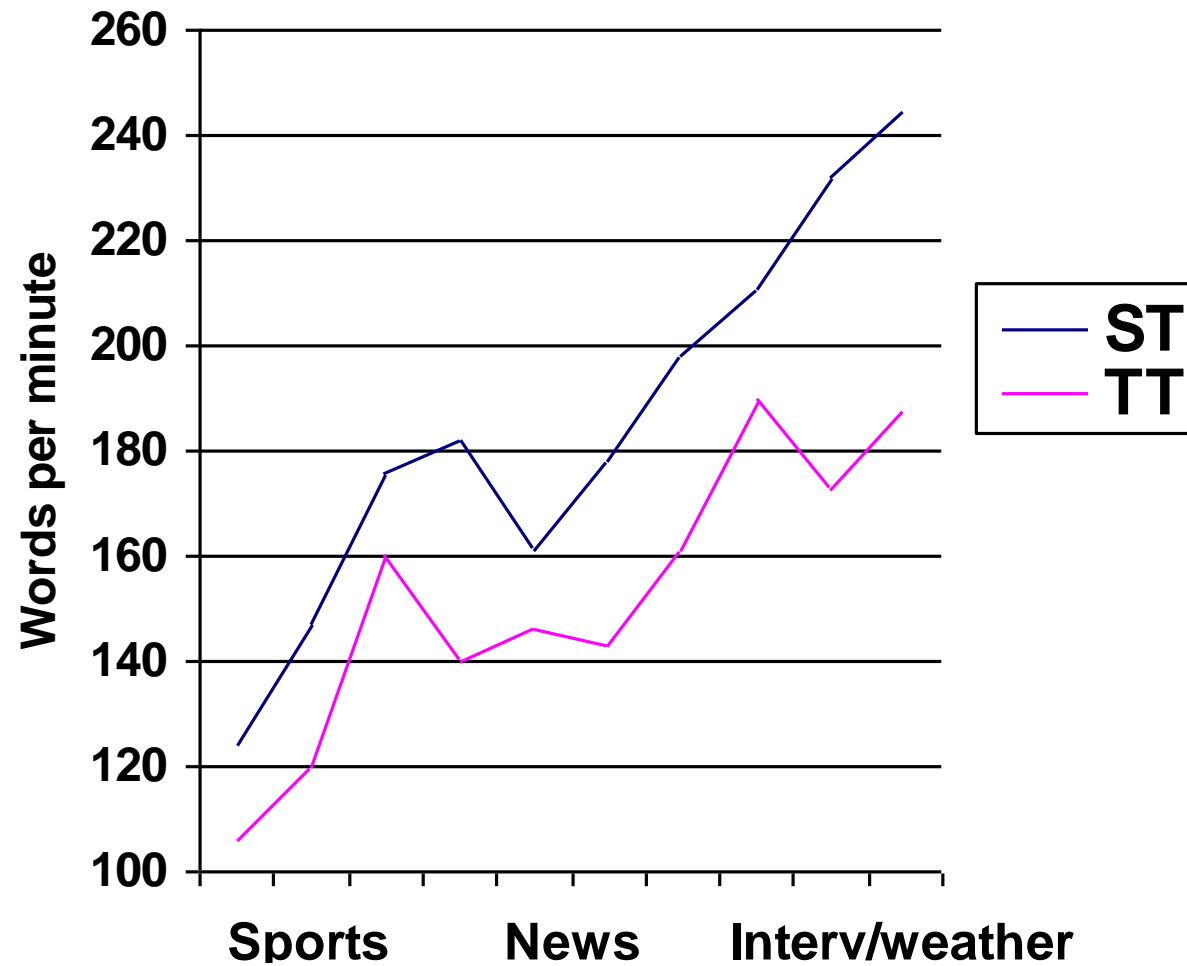
[watch, listen, respk and read]

In **9098**, the financial

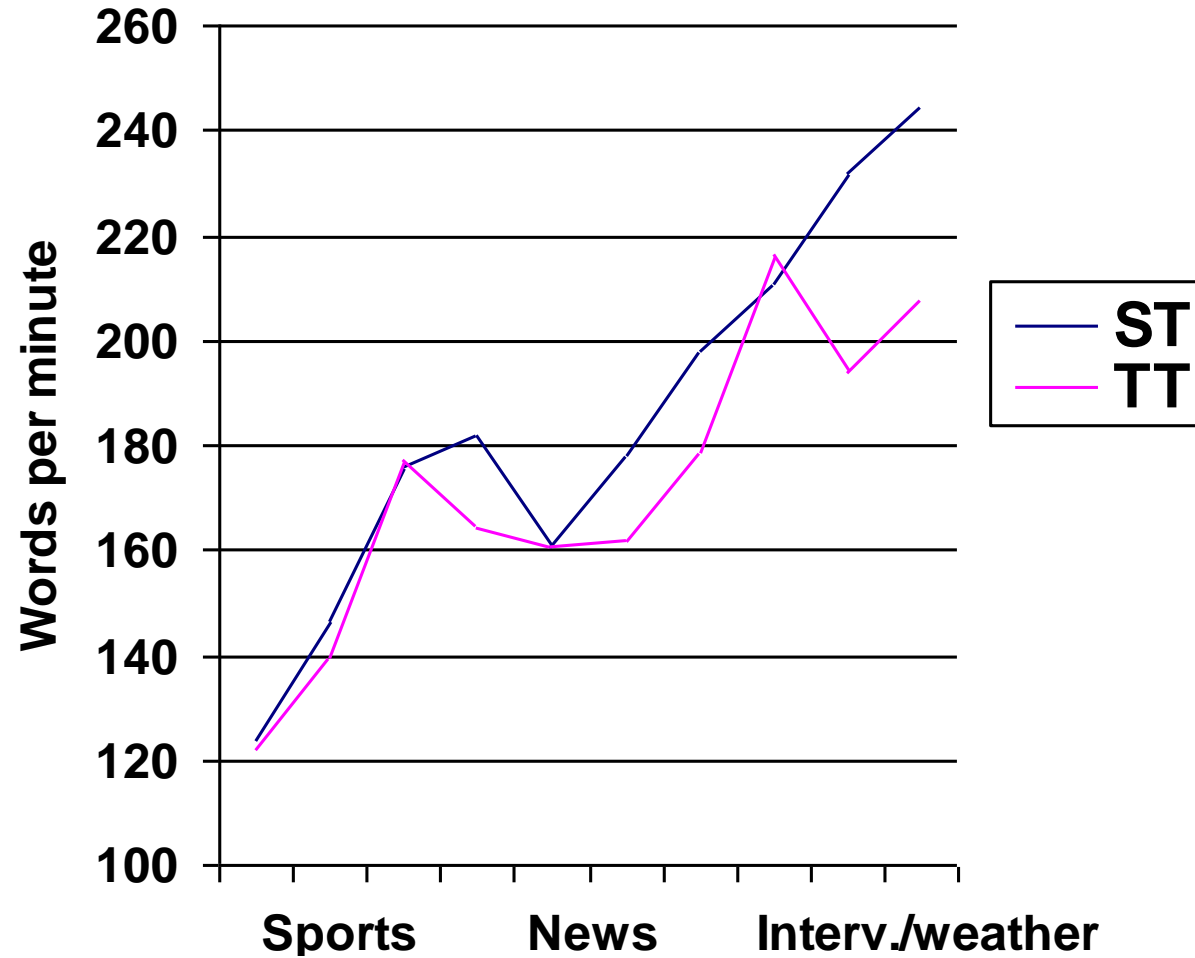
[watch, listen, read and type]

In **1998**, the finan..

Verbatim vs edited



Verbatim vs edited



Impact of verbatim captions on the viewers




Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)


- Realtime voice writing (intralingual respeaking)



Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- 
- Realtime voice writing (intralingual respeaking)
 - Quality in realtime captioning

Quality in Respeaking

Well, you know, you have to try
and put out a good performance,
I mean, yeah, it's kind of a
stepping stone, isn't it, really?

You have to try to put out a good
performance. It's a stepping stone.

The NER model

$$\text{Accuracy} = \frac{N - E - R}{N} \times 100 = \%$$

Correct editions:

Assessment:

Target = 98%

Minor Recognition Errors

What a great goal by **a** Ryan Giggs!

What a great goal **by Ryan** Giggs!

Simon **brown** has been appointed new
chairman of Rolls Royce.

Simon **Brown** has been appointed new
chairman of Rolls Royce.

Standard Recognition Errors

He's a **buy you a bull** asset.

He's a **valuable** asset.

Is it really **attend Tatian**?

Is it really **a temptation**?

Serious Recognition Errors

Public funding for universities has been cut by **50%** this year.

Public funding for universities has been cut by **15%** this year.

Three students **weren't** missing today.

Three students **were** missing today.

The Queen commemorates her birthday on her own.

The Queen commemorates her birthday on heroin.

Standard Recognition Errors

He's a **buy you a bull** asset.

He's a **valuable** asset.

Is it really **attend Tatian**?

Is it really **a temptation**?

It was a documentary about the **path is it**.

It was a documentary about the **Pacific**.

The NER model



The NER model



The NER model




The NER model



Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- 
- Realtime voice writing (intralingual respeaking)
 - Quality in realtime captioning

Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- Realtime voice writing (intralingual respeaking)
 - Quality in realtime captioning
- Interlingual realtime closed captioning (interlingual respeaking)

Contents

Media accessibility and audiovisual translation

Realtime closed captioning (live subtitling)

- Realtime voice writing (intralingual respeaking)
 - Quality in realtime captioning
- Interlingual realtime closed captioning (interlingual respeaking)
 - Training, research and practice

Interlingual realtime voice writing: professional practice

- Interlingual respeaking on TV
- Conferences
- Educational, legal and medical settings

Interlingual realtime voice writing: training

- Mainly in-house (from 12 to 52 weeks, 1-4 hours a week).
- University of Antwerp: face-to-face, English into Dutch
- University of Vienna: face-to-face, English into German
- Universidade de Vigo:
 - online
 - Eng-Spa-Eng
 - 8-week simultaneous interpreting
 - 8-week intralingual respeaking
 - 8-week interlingual respeaking

Interlingual realtime voice writing: Research

- Is interlingual respeaking feasible?
- What is the ideal professional profile?

Contents

Media accessibility and audiovisual translation

Realtime closed captioning / Live subtitling

Realtime voice writing / Intralingual respeaking

Quality in realtime captioning: the NER model

Interlingual realtime closed captioning / Interlingual respeaking

Training, research and practice

Future developments

Future developments

- Automatic realtime voice writing

Automatic vs. manual

- The quality of respoken subtitles is better than that of the automatic ones.
- The quality of some automatic subtitles is impressive:
 - Accurate recognition of structured narration
 - Long stretches of spontaneous speech
- If punctuation is not counted, automatic subtitles are close to the accuracy threshold.
- Too many minor errors (interrupt the flow).
- Too many standard errors in fast and spontaneous speech.

Interlingual realtime voice writing: final questions

- Can/Will this be automatized?
- Can this take off in the US?
- What working conditions would be appropriate?
- Who can do it and how can they be trained here?
- Can the ATA get play a role here?

ATA 60TH ANNUAL CONFERENCE
OCT 23-26 2019 *Palm Springs* CALIFORNIA

Interlingual Real-Time Closed Captions: Where Accessibility Meets Translation



Pablo Romero-Fresco
p.romero-fresco@roehampton.ac.uk
UniversidadeVigo