

# ILSA Multiplier Event

## Interlingual live subtitling with respeaking

# What will be covered in the workshop?

09:00 – 09:20 <b>An Introduction to Respeaking</b>	<ul style="list-style-type: none"><li>• Live subtitling methods</li><li>• What is live subtitling?</li><li>• Current state of affairs</li><li>• Speech recognition software</li></ul>
09:20 – 10:20 <b>Intralingual Respeaking</b>	<ul style="list-style-type: none"><li>• Dictation practice</li><li>• Preparation</li><li>• Intralingual respeaking practice</li><li>• The NER model</li></ul>
10:20-10:35	<b>BREAK</b>
10:35 – 11:45 <b>Interlingual Respeaking</b>	<ul style="list-style-type: none"><li>• Current state of interlingual respeaking</li><li>• Interlingual respeaking practice</li><li>• The NTR model</li><li>• Research in interlingual respeaking</li></ul>
11:45 – 12:00 <b>Closing Discussion</b>	<ul style="list-style-type: none"><li>• Skills, challenges and the future of ILS</li></ul>

# AN INTRODUCTION TO RESPEAKING

# Live subtitling methods



# Comparison of live subtitling

(Lambourne, 2007)

	<b>Velo- type</b>	<b>Dual Keyboard</b>	<b>Steno</b>	<b>Respeak.</b>
<b>Delay</b>	Mid	Mid	Low	Low
<b>Speed</b>	Mid	Mid/High	Very high	Mid/High
<b>Accuracy</b>	Mid	Mid/High	High	High
<b>Difficulty</b>	Mid	Mid	High	Mid
<b>Cost</b>	Mid	Mid	Mid/High	Low
<b>Errors</b>	Mid	Low	Mid	Mid

# What is live subtitling?

## RESPEAKING (or REAL-TIME 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)

## INTERLINGUAL RESPEAKING

“With regard to the process, ‘interlingual respeaking’ [...] is really a form of simultaneous interpreting, while the product, in the intralingual as well as the interlingual variant, is a set of subtitles” (Romero-Fresco & Pöchhacker 2017: 158)

# Respeaking in action



<https://www.youtube.com/watch?v=fJ2waoxrJ2c>

# Origins of respeaking



# Current state of affairs

- Respeaking has overtaken stenography as the preferred form of live subtitling.
- North America and Italy mainly use stenography.
- Australia and the UK use a mix, although respeaking is more common in the UK.
- Spain, Switzerland, Belgium and France resort almost exclusively to respeakers, at least for TV subtitles.

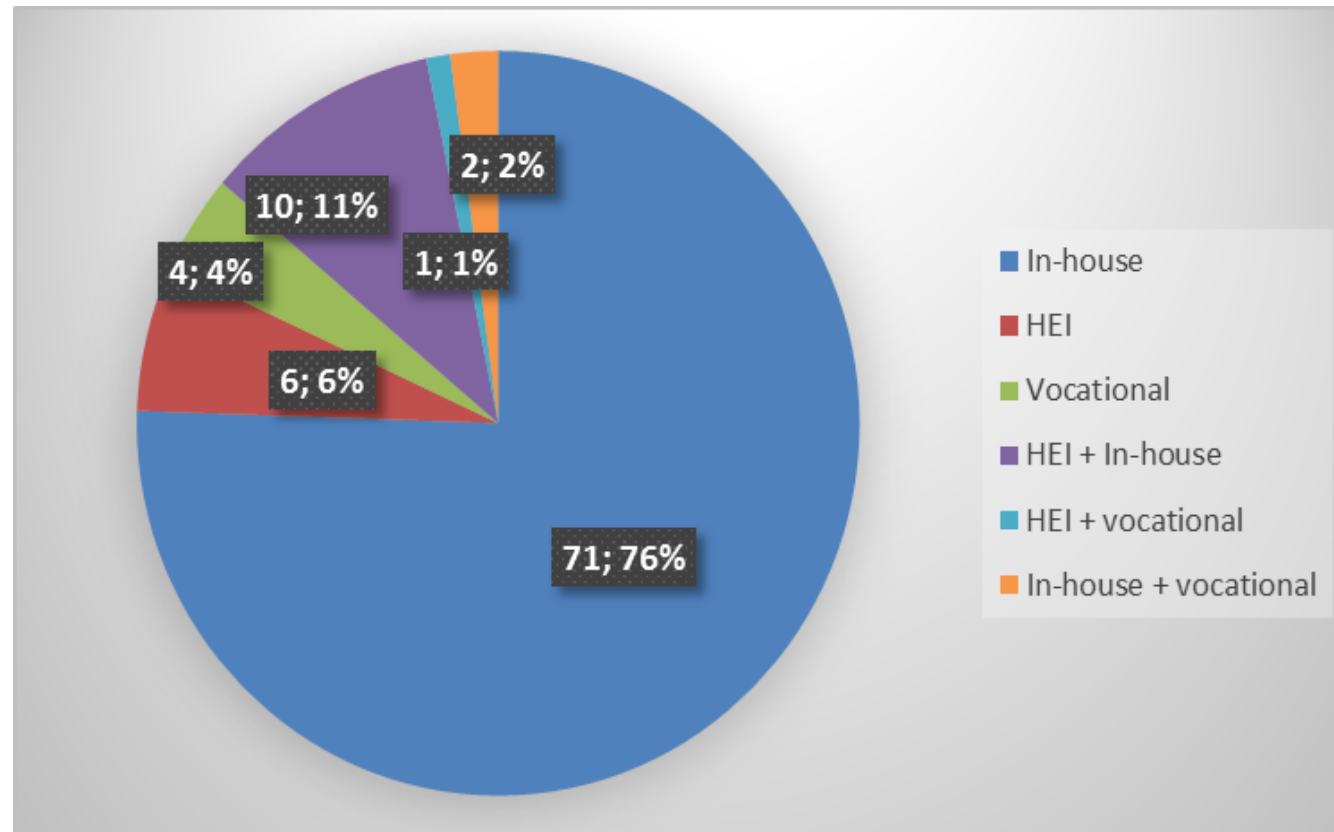


# Respeaking as a profession (ILSA IO1 results)

- Highly educated women under 40 with a BA and language-related MA.
- Working as respeakers for 6 years and typically work for subtitling companies (in-house or freelance).
- Demand for quality has changed how respeakers work = scripts.
- Some differences in equipment depending on country
- Accuracy and delay is approached differently depending on country.
- Respeakers do not normally respeak for more than one hour.

# Respeaking training

- Respeaking was introduced in Europe in 2001, but formal training did not start until 2007.
- Due to lack of research, codes of practice and guidelines, companies trained their own respeakers.



# Respeaking training

- Respeaking training at university level is limited.
- Respeaking 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: Respeaking is to subtitling what interpreting is to translation (a leap from the written to the oral without the safety net provided by time).

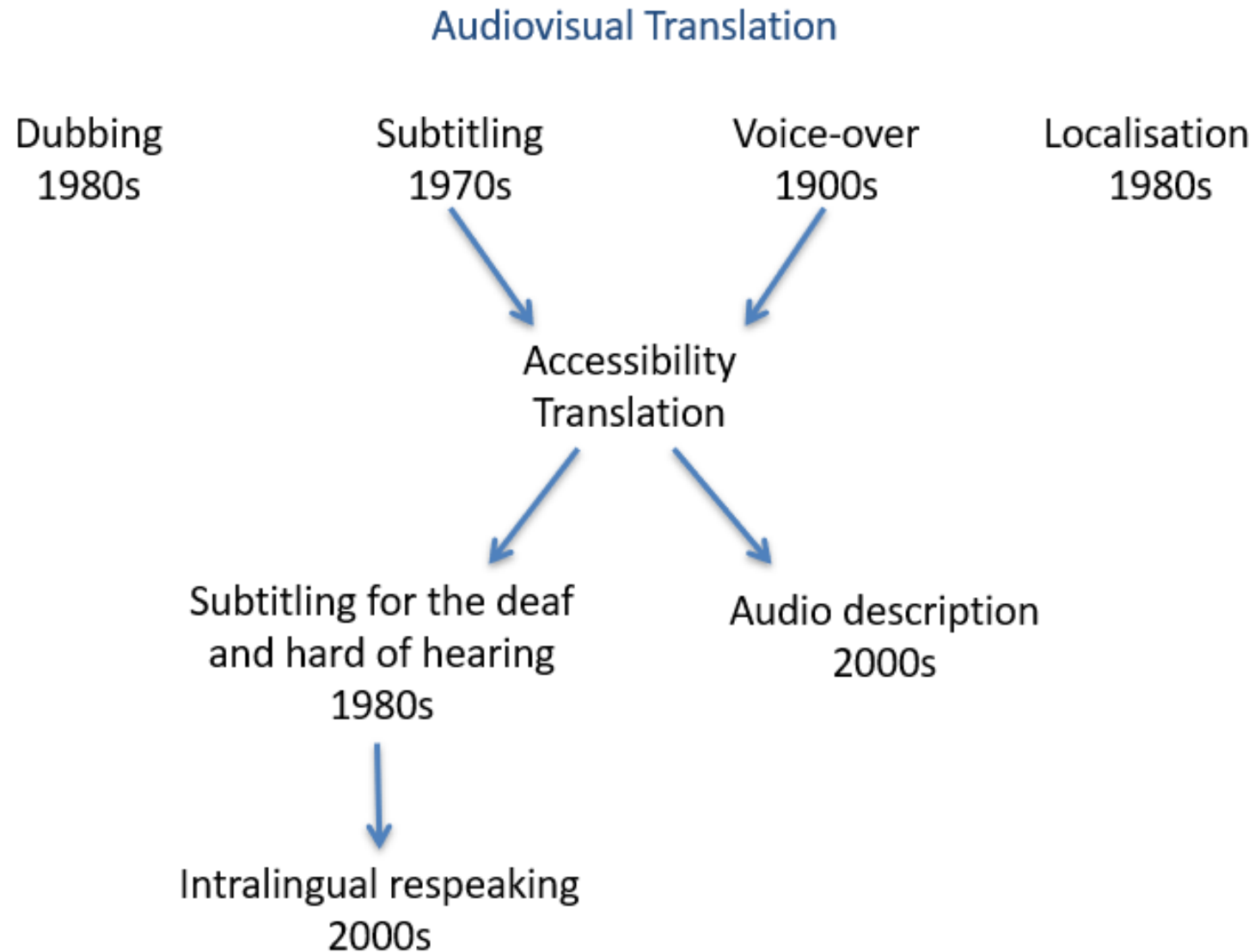
Process: interpreting

Result: subtitling

Technology: speech recognition

(Romero-Fresco, 2011)

# Where does intralingual respeaking sit?



# Respeaking training

- Length of training and course structures differ.
- Recurrent elements include:
  - Focus on dictation and working with software.
  - Awareness of subtitling, SDH, DHOH audience, subtitling software and conventions.
  - Interpreting skills, such as multitasking, listening, comprehension, reformulation.

# Reoccurring elements in respeaking training

Theoretical  
introduction

Software  
Voice profile  
Dictation practice

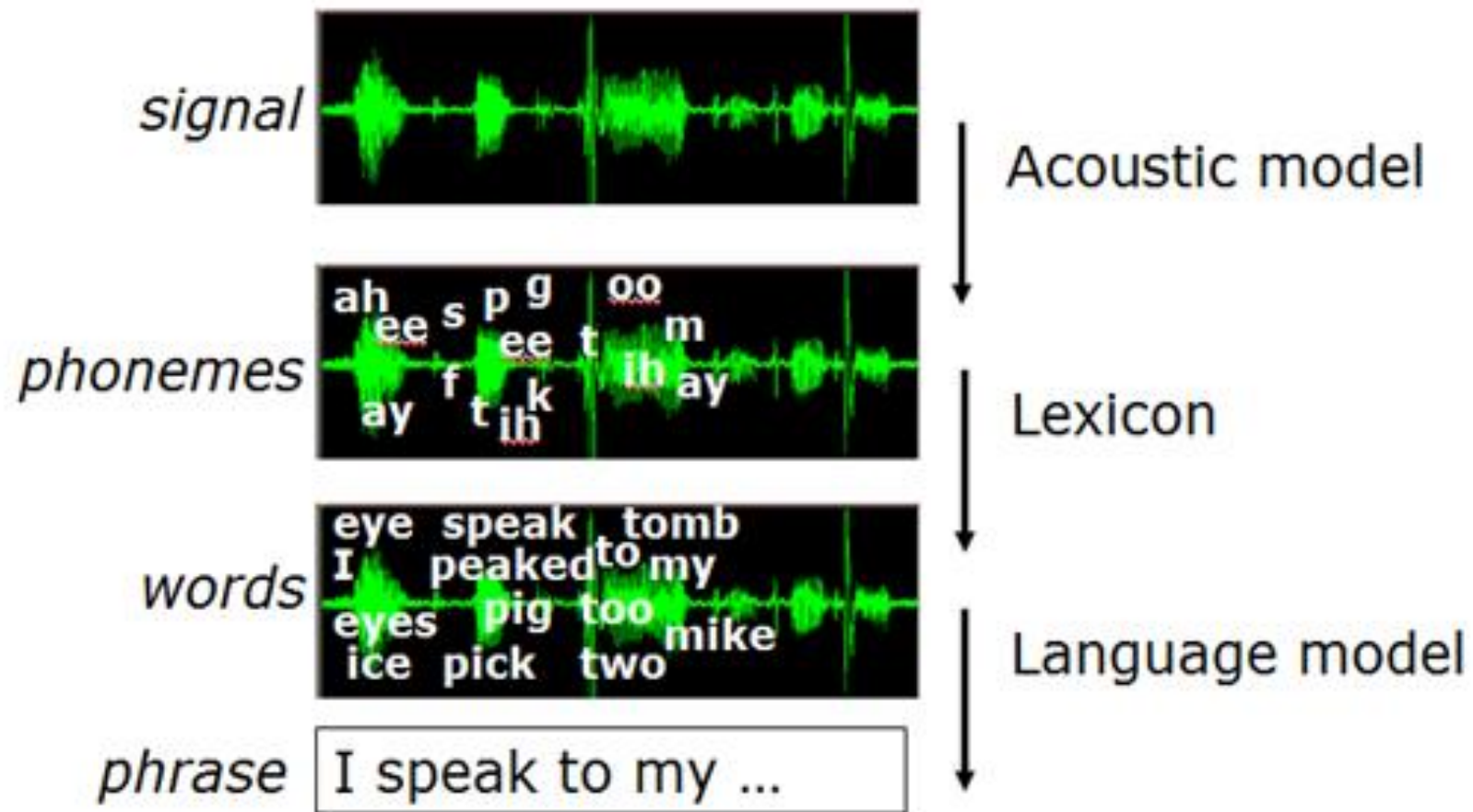
Respeaking practice





# How does speech recognition work?

- Speaker dependent vs. speaker independent



# General knowledge of speech recognition

In + 2009 + the + financial + recession

In + **2009** + the + financial + recession + brought

In + 2009 + **the** + financial + recession + brought + about

In + 2009 + the + **financial** + recession + brought + about + the +

In + 2009 + the + financial + **recession** + brought + about + the + loss

# INTRALINGUAL RESPEAKING DICTATION

# Setting up your voice profile

## Microphone

1. Ensure your microphone is adjusted comfortably and is in a stable position.
2. The microphone should be at either side of your mouth and at a distance of 2cm.

## Create a user profile

1. Select the 'short training' option.
2. You should read the text as you plan to dictate.
3. If a word is not recognised it will be indicated and you have to read the word again. It is advisable to repeat the last few words before the unrecognised word too.
4. Dragon will offer to analyse your documents and emails, please skip this step.

## Speed Settings

1. Adjust the speed setting so the words appear faster on screen.
2. DragonBar > Tools > Options > Miscellaneous.
3. In Speed vs. Accuracy set to 'fastest response'.

# Tips for dictation practice

- Use a **steady pace**
- Use **short and concise sentences**
- Remember to **enunciate punctuation**
- You must **briefly pause for the text to be released on screen**. An ideal time to pause would be after each sentence, this way you will remember to pause each time you dictate 'full stop'.
- You should **speak clearly** but without raising or lowering your voice. It is ideal if your voice is **slightly robotic** to facilitate the software's recognition.
- If the software misrecognises a word do not continue to dictate the word in the same way.  
Remember that **you and the software form a team** and must **work together** so try to dictate the word differently.

# Tips for dictation practice

Dragon hears everything!

You must **avoid**:

- Challenging the software
- Laughing at the mistakes
- Commenting to your neighbour

# INTRALINGUAL RESPEAKING DICTATION PRACTICE

# EXERCISE 1

## Dictation Practice

Open the folder on your desktop titled 'Workshop Exercises' and complete exercise 1.



# EXERCISE 1

## Dictation Practice

Use some of the texts in exercise 1 to dictate at a comfortable pace. Please calculate your speed in words per minute to identify your ideal dictation speed.

Now, dictate the text again at a faster pace in order to identify your maximum dictation speed.

# Intralingual respeaking in action

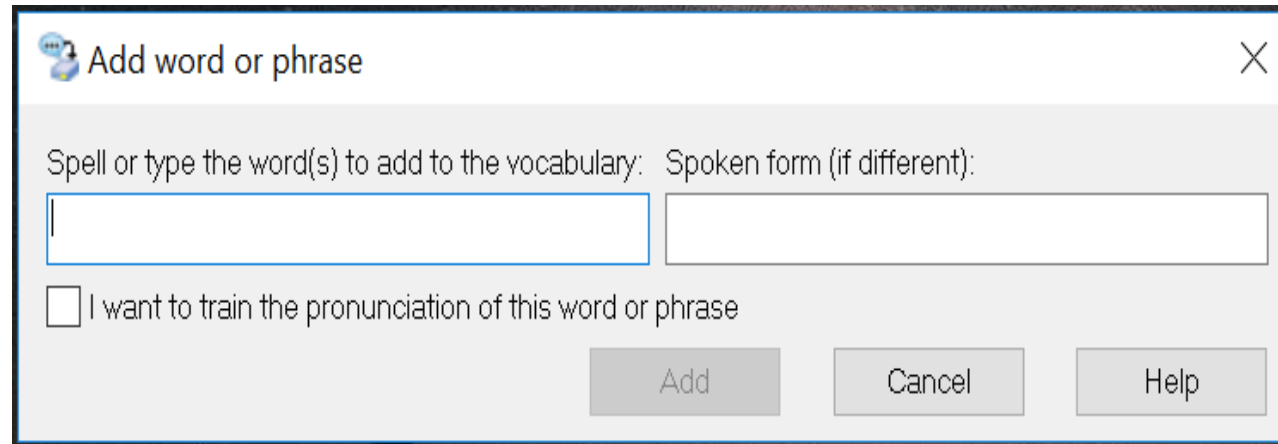
# INTRALINGUAL RESPEAKING PREPARATION

# Word lists

- Interpreters make and learn their glossaries, respeakings only make them. The software does the learning.

# Adding words or phrases

- To add a word to Dragon go to Vocabulary > Add new words or phrase.
- Type the word in and select the option to train the word by dictating it to Dragon.



The screenshot shows a dialog box titled "Add word or phrase" with a close button (X) in the top right corner. Inside the dialog, there are two text input fields. The first field is labeled "Spell or type the word(s) to add to the vocabulary:" and the second field is labeled "Spoken form (if different):". Below these fields is a checkbox labeled "I want to train the pronunciation of this word or phrase". At the bottom of the dialog, there are three buttons: "Add", "Cancel", and "Help".

# Macros

- Respeakers can create macros. Macros are vocal shortcuts or customised commands.
- Respeakers use them to “ensure that certain labels or phrases are always correctly formatted, and save a great deal of time and effort” (Marsh 2006:32).

# Creating macros

“word + macro” or “mac + word”

“cheer macro” or  
“maccheer”

=

CHEERING AND  
APPLAUSE

“whistle macro” or  
“macwhistle”

=

WHISTLE BLOWS

“mobile macro” or  
“macmobile”

=

MOBILE RINGS

# INTRALINGUAL RESPEAKING PRACTICE



# Intralingual Respeaking

## EXERCISE 2 - Practice

Respeak the video “EXERCISE 2a\_Desperate Housewives Narration” from English into English.

In order to respeak the videos “EXERCISE 2b – Barak Obama Speech” and “EXERCISE 2c - Michelle Obama Speech”, check the word docs with the same titles to add some key words to Dragon.

# INTRALINGUAL RESPEAKING UNDERSTANDING AND MANAGING ERRORS

# The NER model

(Romero Fresco & Martínez, 2015)

**N = Number of words** in respoken text, including commands

**E = Edition errors** (deletions, substitutions, insertions)

- Minor (-0.25)
- Standard (-0.5)
- Serious (-1)

**R = Recognition errors** (caused by human or software)

- Minor (-0.25)
- Standard (-0.5)
- Serious (-1)

**CE = Correct Editions** (correctly edited text)

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

CE:

Comments:

# 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.

**For** people are still missing following Sunday's tornado.

**Four** people are still missing following Sunday's tornado.

# 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**.

# 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.

He never talks **dirty**.

He never talks **to Rudy**.

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

Three students **were** missing today.

# Idea units

**Dependent idea units** = pieces of information i.e. the 'who', 'what', 'where', 'when', 'why'

**Independent idea units** = a full sentence or a full message

# Idea units

**Independent idea unit**

The two victims of a fatal fire in Melbourne are yet to be formally identified.

The blaze started this morning at the front of the house.

**Dependent idea unit**

**Dependent idea unit**

**Independent idea unit**



# Error severity

**Minor (-0.25)** – Errors that affect dependent idea units or can be recognised.

**Standard (-0.5)** – Errors that affect full sentences, are confusing or cause loss of information.

**Serious (-1)** – Errors that introduce false or misleading information.

# Minor Edition Errors

Tomorrow, I have a meeting at 1pm.

Tomorrow, I have a meeting ~~at 1pm~~.

# Standard Edition Errors

Your job today is to cook one classic dish  
with your stamp on it.

Omitted

# Serious Edition Errors

Hillary Clinton gave a speech in **Beijing**.

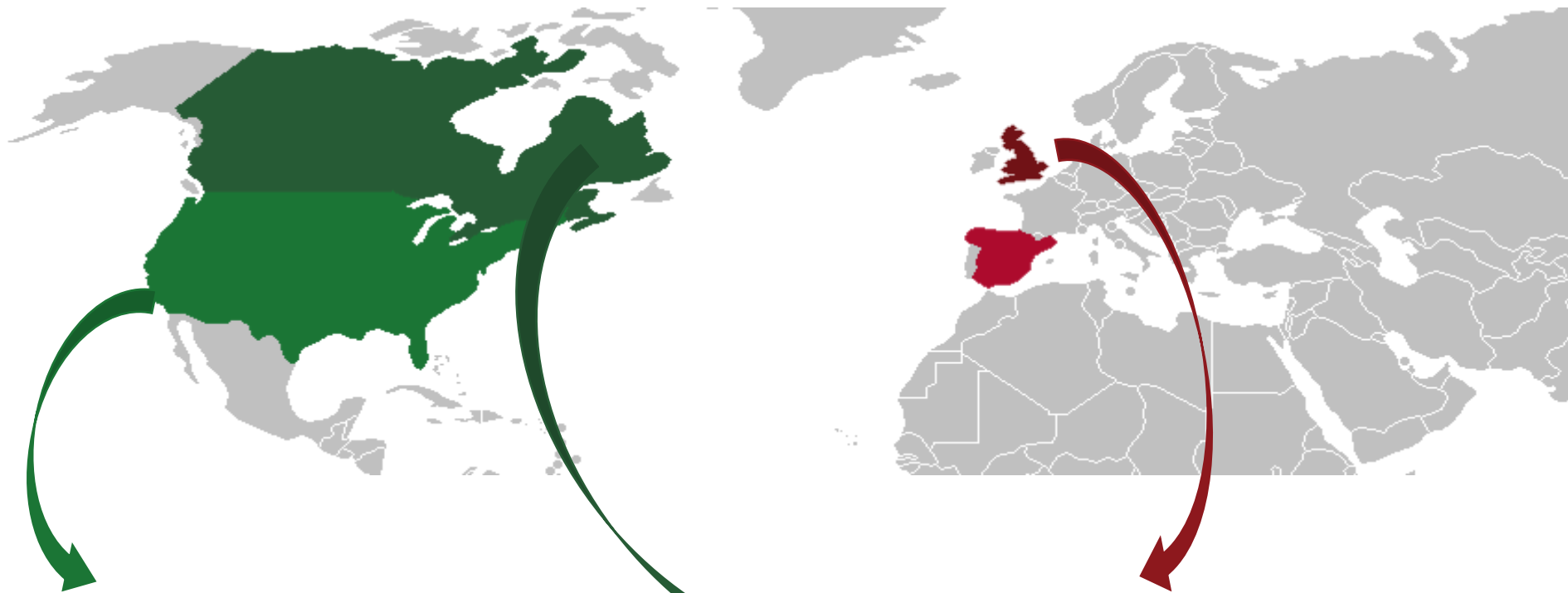
Hillary Clinton gave a speech in **Berlin**.

The total cost was **£1,000**.

The total cost was **£2,000**.

# Respeaking regulations in different countries

I



- Non-metric (soft) approach:
  - No official quality assessment
  - Based on user complaints
  - AR calculations based on the number of *words* in CC.
- Canada has just adopted NER
- The US may change approach soon

- Metric (hard) approach:
  - Official quality assessment
  - NER model
  - AR calculations based on the number of *ideas* in subs.
- First adopted in the UK
- Also used in Spain + many others

# INTRALINGUAL RESPEAKING PRACTICE

# EXERCISE 3

## Practice

Consult 'Exercise 3' in your exercises folder.

Respeak 'Emma Watson speech', 'Virtual Reality news' and/or 'DiCaprio, Pitt and Robbie\_Once Upon a Time ... in Hollywood' from English into English.



Break Time!



# INTERLINGUAL RESPEAKING

# Current state of interlingual respeaking (research and training)

- A handful of studies have explored quality assessment, ear-voice span and pauses, crisis points in respeaking (Robert & Remael, 2017; Romero-Fresco & Pochhacker, 2017; Chimel et al., 2017; Szarkowska, Dutka & Pilipczuk, 2017).
- Training in ILS is scarce (MA level in Interpreting and Translation degrees).
- Duration: ranges from 12 to 52 weeks (1-4 hours).
- Short training: dictation, respeaking practice, software knowledge.
- Long training : theoretical introduction, company visits.
- Quality assessment methods are unknown.

# Current state of interlingual respeaking (professional practice)

- Interlingual respeaking on TV: Wales, Belgium, The Netherlands
- Conferences
- Educational, legal and medical settings
- Research events: multiplier events.

# Current state of interlingual respeaking (professional practice – working conditions)

- To be discussed at the end of the workshop

# INTERLINGUAL RESPEAKING PRACTICE

# EXERCISE 4

## Practice

Take 5 minutes to search for a video in your native language to respeak into English.

The video should be similar to those we used for intralingual respeaking exercises.

Try to pay particular attention to recognition errors.

- What do you think has caused these errors?
- How could they be avoided in the future?

- Try searching the EU speech repository for speeches of varying difficulty in your native language.

- We will complete three interlingual exercises. Try to use videos of varying difficulty:

-**Easy:** narration, speech (low speech rate)

-**Intermediate:** news, weather, sports (medium speech rate)

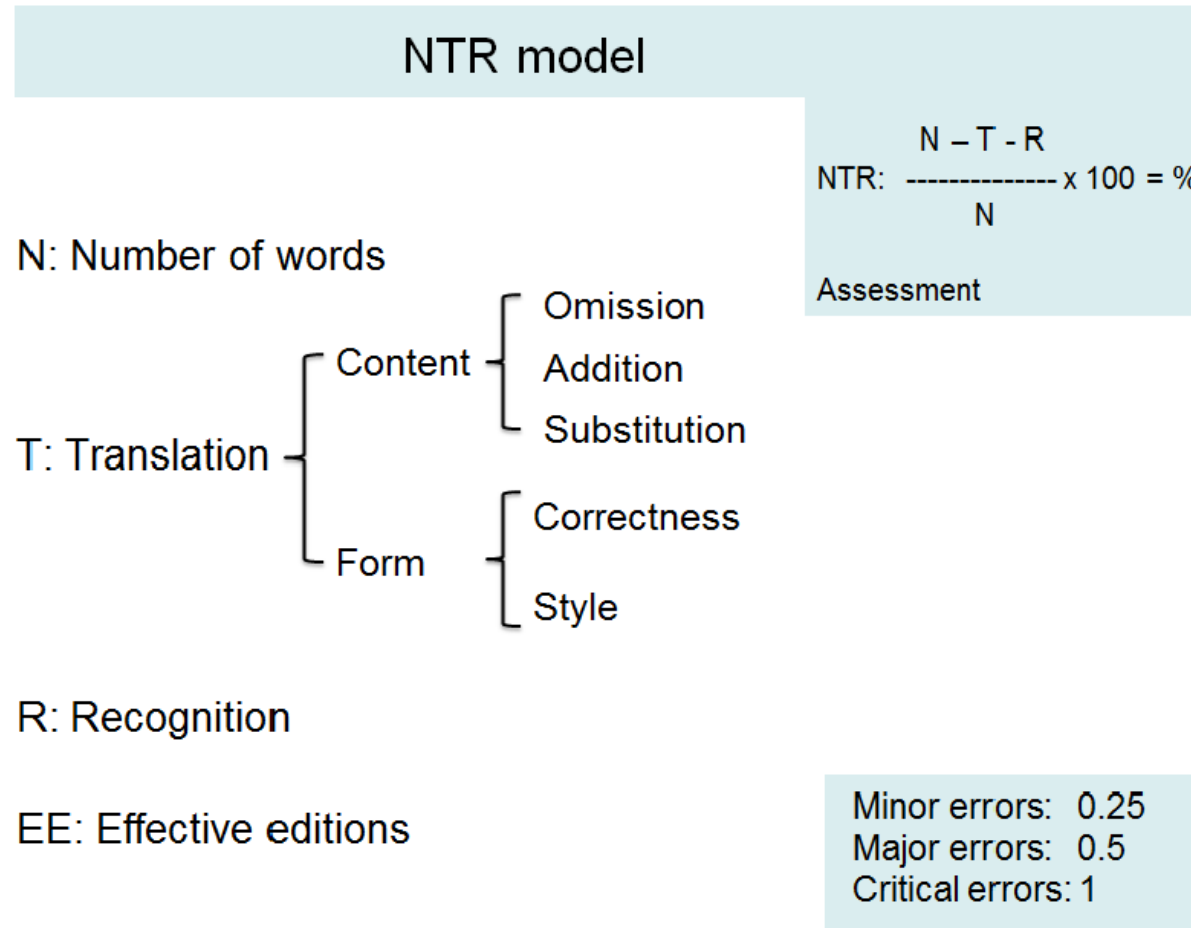
-**Advanced:** chat show, interview, multiple speakers (high speech rate)

# INTERLINGUAL RESPEAKING UNDERSTANDING AND MANAGING ERRORS



# The NTR model

(Romero-Fresco & Pöchhacker, 2017)



# Minor Translation Errors

The smart beer tap will not replace bar staff, it will simply be there to help them.

Los inventores dicen que el tirador no reemplazará a los camareros ~~estará para~~ **ayudarles**.

I started questioning gender-based assumptions a long time ago.

Empecé a **asumir (cuestionarme)** algunos roles de género hace mucho tiempo.

# Major Translation Errors

Put this together with the physical effort of working in the garden, and we see significant improvements in many mental health disorders such as depression as I mentioned.

~~Esto junto al esfuerzo físico de trabajar en el jardín da mejoras significativas en varios trastornos mentales como la depresión.~~

Para perder peso, por ejemplo, se puede ir al gimnasio, pero algunas personas no nos gusta mucho el gimnasio.

# Critical Translation Errors

But pouring a pint isn't always easy, it can take a bit of practice.

Servir cerveza **es fácil** aunque hace falta práctica.

I was confused about being called bossy, because I wanted to direct the plays.

Me llamaban mandona porque quería dirigir los **juegos (teatros)**.

# Example of NTR analysis

Original text (transcribed audio)	Respeaking-based subtitles	Errors
Ladies and gentlemen, I live in Istanbul, which as you probably know is a huge metropolis and in fact one of the biggest cities in the world. As such, it's not a city that is particularly well known for its green spaces and indeed it is a very stressful place to live. However, I count myself very lucky because I happen to have a garden of my own where I can plant flowers and even vegetables, which I then enjoy eating. Now I find that after spending half an hour or an hour or so in the garden I feel much happier and much more relaxed. So any time I'm feeling a bit down, I've got a bit of the blues what I do is I go out and I spend some time digging up some weeds or planting some flowers and this makes me feel much better.	Señores y señoras, vivo en Estambul, una enorme ciudad, de hecho, una de las mayores del mundo. Por lo tanto, no es una ciudad especialmente conocida por sus espacios verdes, de hecho, puede ser un lugar muy estresante para vivir. Sin embargo, me considero muy afortunada, porque tengo un jardín propio, donde puedo plantar flores, e incluso verduras que me encanta comer. Según yo lo veo, después de pasar media hora, o así en el jardín, me siento <b>más contenta y</b> (1) mucho más relajada. Por lo tanto, cada vez que me siento algo deprimida lo que hago es salir y pasar algo de tiempo <b>de Servando (desherbando)</b> (2) o plantando flores. Todo esto me hace sentir mucho mejor.	1.MinT (cont-omiss) (0.25): The viewer would miss some of this description.  2. MajR (0.5): The error cannot be recognised.  EE: No relevant information is lost here.  3. MinR (0.25): The error can be recognised.  4. MinT (cont-omiss) (0.25): Omission of a dependent idea unit.  5. MinR (0.25): The error can be recognised.

Figure 3: Example of NTR analysis for a test

# Example of NTR analysis

Accuracy rate	
MinT: 8 (8 x 0.25 = 2) (cont-omiss) x 6 (cont-add) x 1 (form-corr) x 1  MajT: 0  CritT: 0  Total: 2	MinR: 11 (11 x 0.25 = 2.75)   MajR: 1 (1 x 0.5 = 0.5)  CritR: 0  Total: 3.25
NTR accuracy rate N = 692 (622 + 70) $\frac{692 - 2 - 3.25}{692} \times 100 = 99.24\% (8/10)$	
EE: 4	
Assessment	
<p>The accuracy rate reaches 98%.</p> <p>The overall quality of the respoken text is excellent. The majority of minor translation errors (8) refer to omissions of dependent idea units (6), which the respeaker could have used as a strategy to keep up with the text. The recognition is acceptable as minor errors (11) are all recognisable and only 1 error causes loss of information. The text is coherent and very easy to follow.</p>	

# INTERLINGUAL RESPEAKING IO2 RESULTS OF THE ILSA PROJECT

# The short online course

## WEEK 1

Pre-experiment  
questionnaire

Setting up Dragon  
and creating a user  
voice profile

Dictation practice

## WEEK 2

Guidelines for  
intralingual  
respeaking

Subtitling readings

Intralingual  
respeaking practice x  
2

## WEEK 3

Guidelines for  
interlingual  
respeaking

Interpreting  
readings

Interlingual  
respeaking practice  
x 2

## WEEK 4

Test x 2

Post-experiment  
questionnaire



# IO2 Results

- Online course taught to 44 participants
- Dictation, intralingual and interlingual practice
- Most popular errors: omissions and mistranslations
- Essential skills: multitasking, live translation, dictation, language, communication

GROUP	OVERALL	TEST
Subtitlers	97.32% (3/10)	97.59% (4/10)
Interpreters	97.40% (3.5/10)	97.62% (4/10)
AVERAGE:	97.36% (3/10)	97.61% (4/10)

# Summary of IO2 data

- ILS seems feasible (97.6%, 4/10).
- Interpreters perform better.
- There is a greater difference between 'good' and 'poor' performers than between groups.
- Translation and recognition are equally important and challenging.
- Good performers have around 50% fewer T and R errors than bad performers, including consistently less serious errors.
- Bad performers struggle to keep up and as a result omit too many full sentences, mistranslate the source text and dictate less clearly.
- Those who don't have a clear-cut interpreting background seem to struggle trying to keep up with the text, as a result they have more omissions, more mistranslations and more recognition errors.

# EXERCISE 5

## Practice

Take 5 minutes to search for another video in your native language to respeak into English.

This time, pay extra attention to translation errors.

- Which sub-types of translation error did you make?
- Why do you think they occurred?

# INTERLINGUAL RESPEAKING

RESULTS OF THE FIRST INTERLINGUAL RESPEAKING  
COURSE, UNIVERSITY OF VIGO

## **AN INTRO TO INTERLINGUAL RESPEAKING**

- Producing a live translation
- Recognition errors
- NTR model

## **SPEECHES**

Exercise: Omissions

Reading: Tips for  
effective interlingual  
respeaking

## **SPORT**

Exercise:  
Mistranslations

Reading: The pilot  
study

## **NEWS**

Exercise: Recognition

Reading: The SMART  
study

## **WEATHER**

Exercise: Sharing  
strategies and  
calculating delay

Reading: Live  
subtitlers: who are  
they?

## **ENTERTAINMENT**

Exercise: Form errors

Reading: Comparisons  
of interpreters and  
subtitlers

## **CHAT SHOWS**

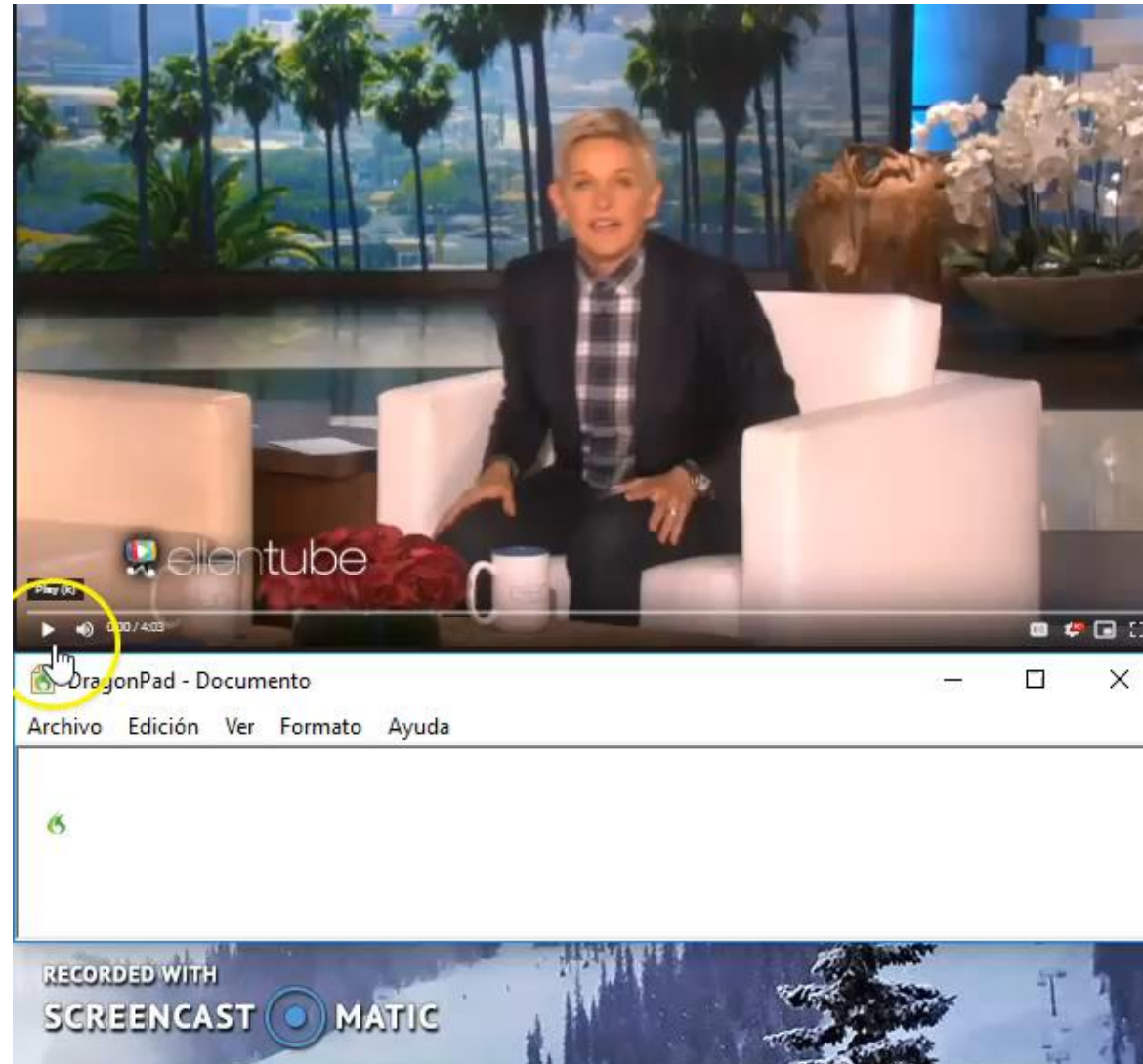
Exercise: Dictation and  
pace review

Reading: A  
competence-oriented  
task analysis of  
interlingual respeaking

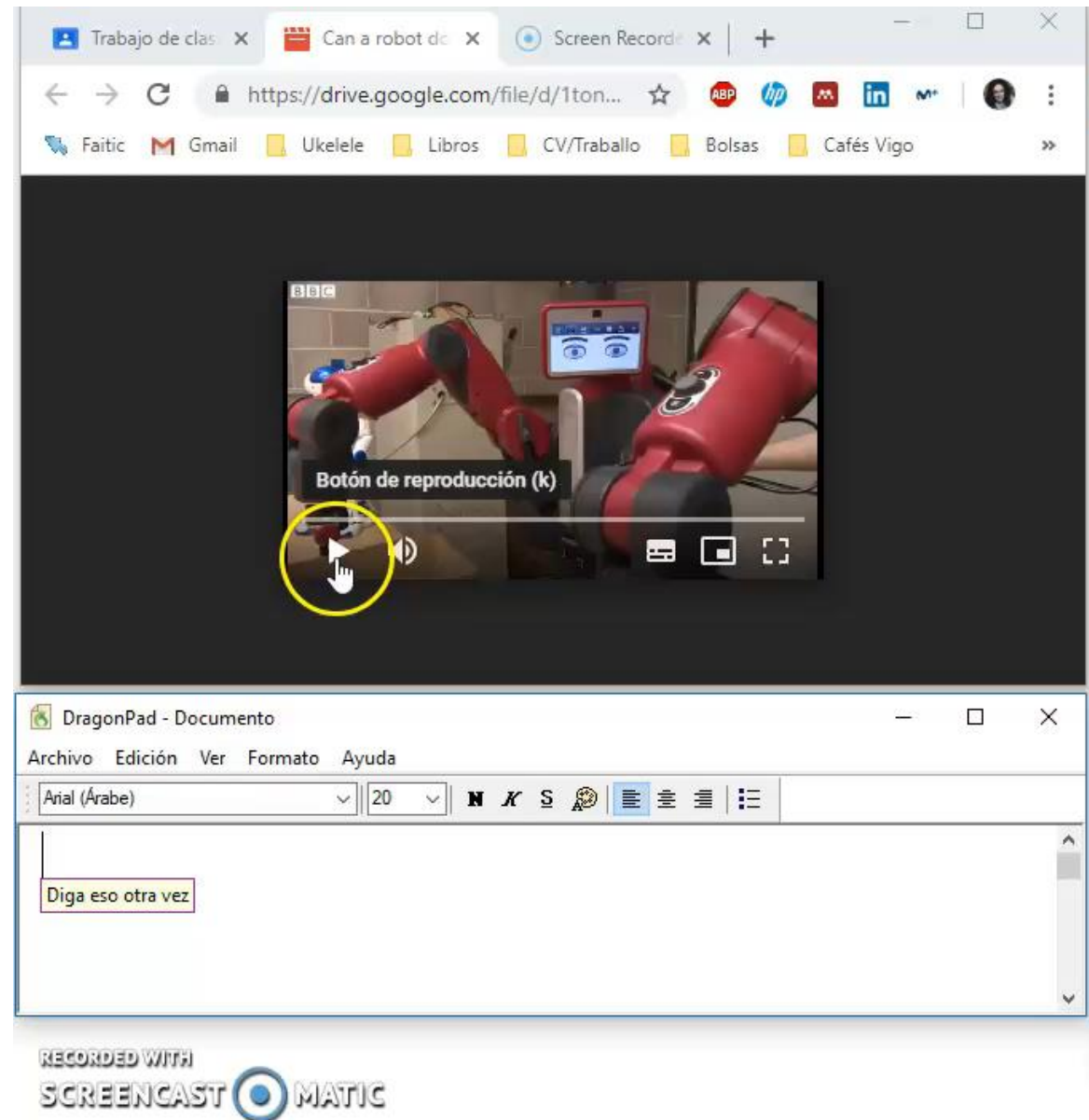
## **FUTURE OF RESPEAKING**

- Full NTR analysis
- Discussion of  
difficulties
- Tips for new  
respeakers
- Course feedback

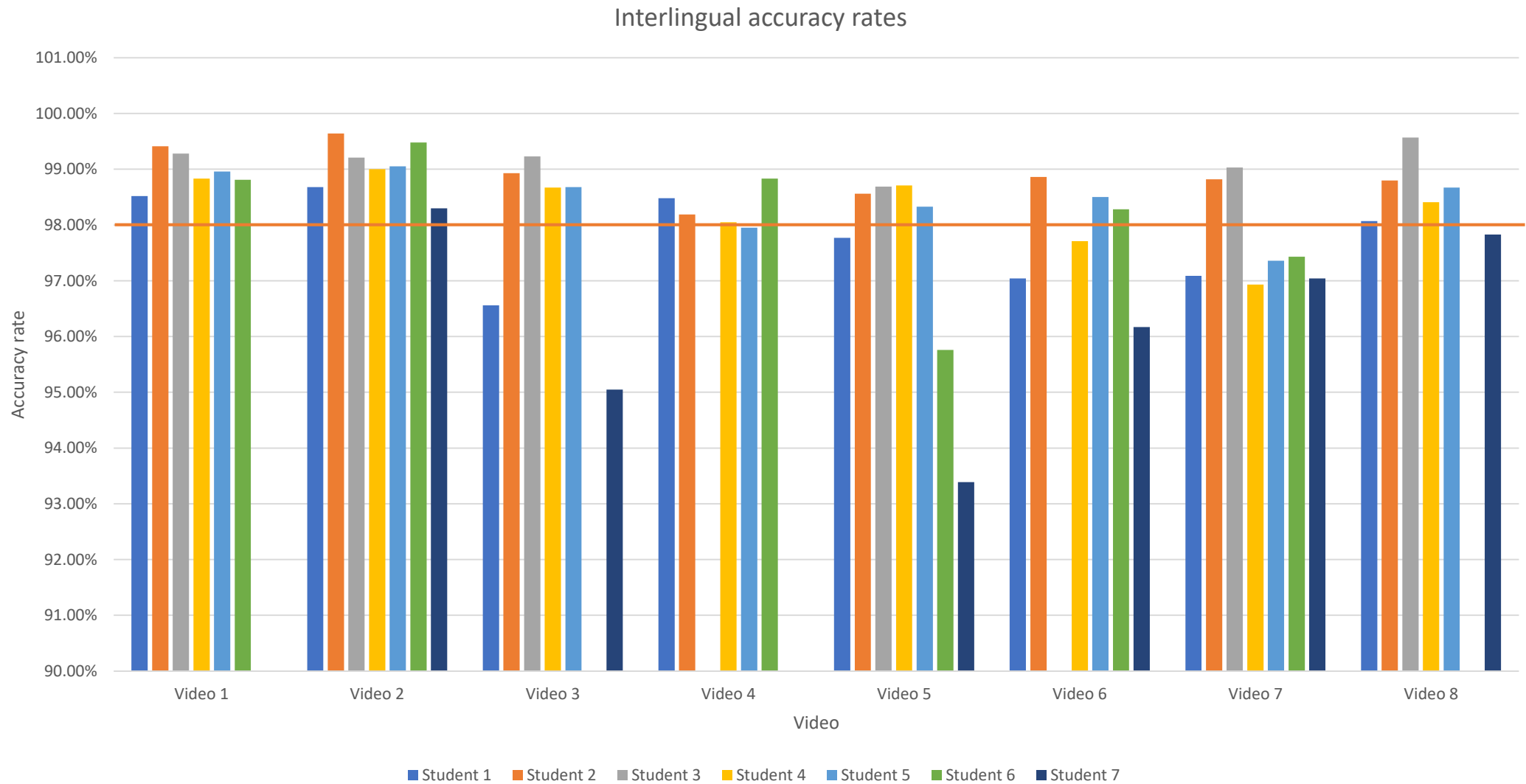
# Interlingual Respeaking in action



# Interlingual Respeaking in action



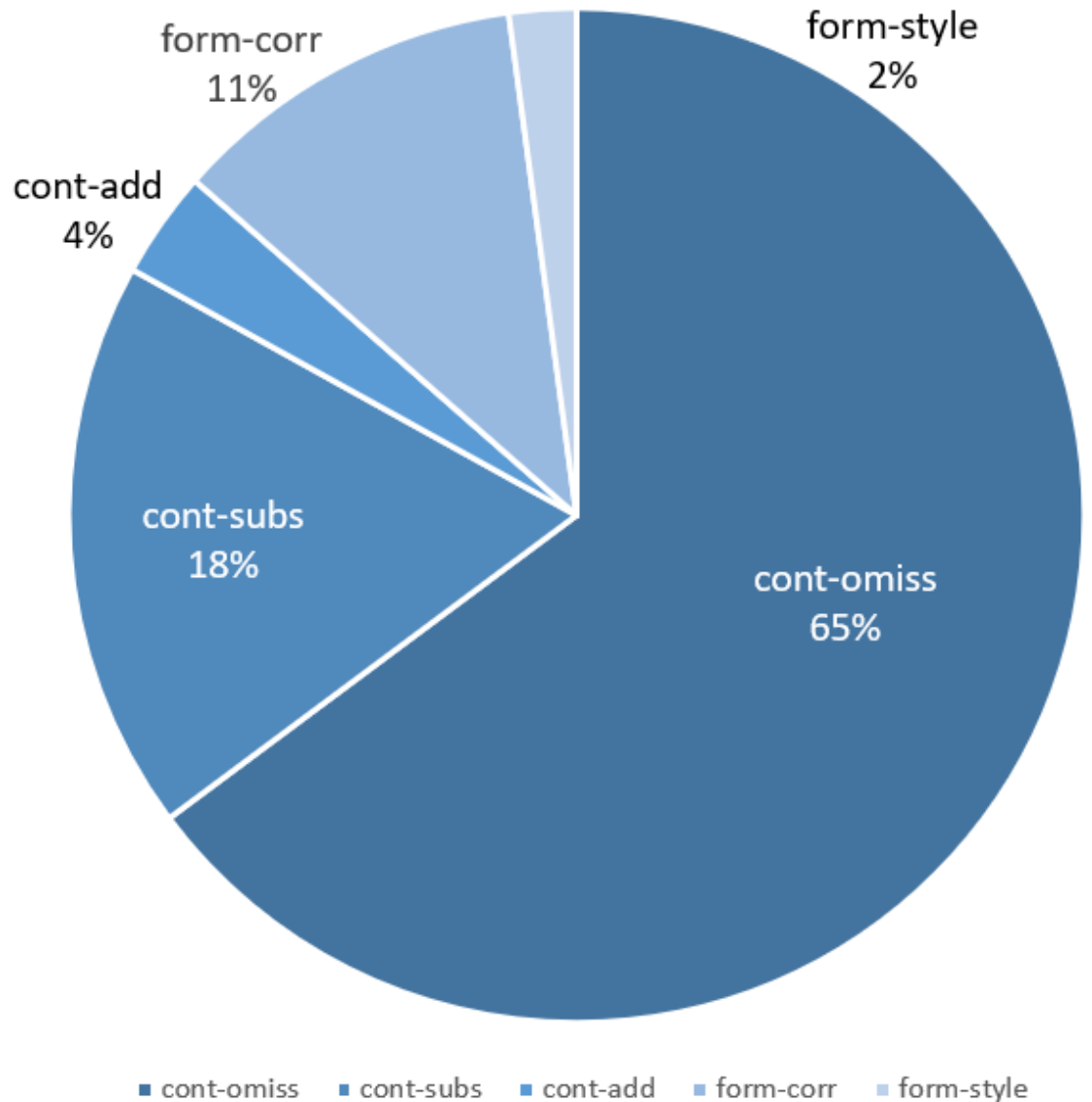
# Vigo course data





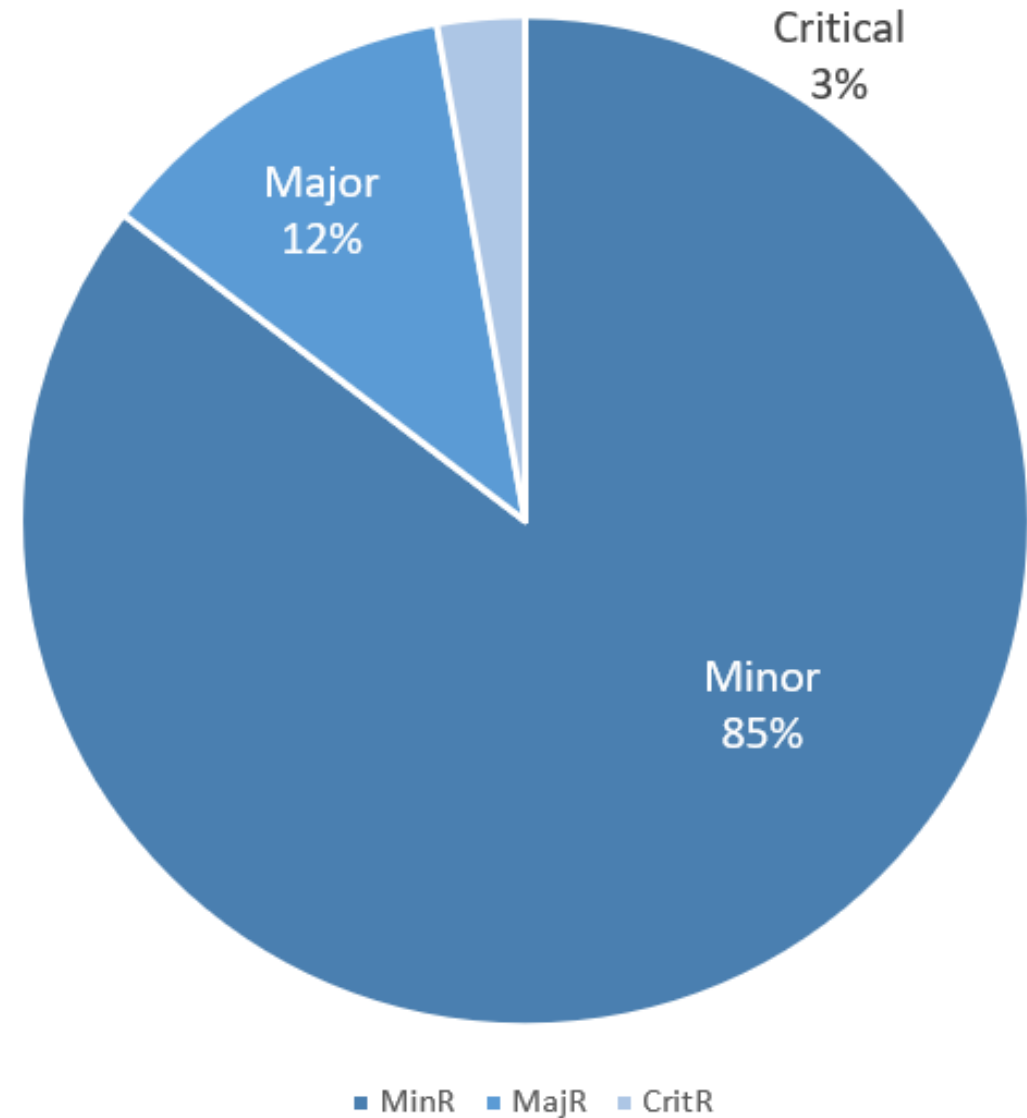
# Translation errors

- Most common = omissions and mistranslations.
- High speech rate = more omissions.
- Low speech rate = least omissions.
- Omissions could be due to the genre, content and delivery of the video.



# Recognition errors

- High speech rate = more R errors (video 7).
- No further correlation between R errors and speech rate.
- Errors vary - 12.5 errors per video to 1.1.
- One student made a total of 9 errors across 8 videos and made no errors in 3 videos.



# Summary of Vigo course data

- Overall accuracy rate = 98.14%.
- Omissions and substitutions are the most common T errors.
- Excellent recognition can be achieved with extensive voice profile training.
- The complexity of videos (content, terminology, speed, delivery) causes difficulty for live translation.
- High speech rates can cause more omissions, but not necessarily more recognition errors.
- In the main experiment, after three weeks of training, participants did not perform well in the speech. In the Vigo course, students achieved good results in a similar speech.

# EXERCISE 6

## Practice

Take 5 minutes to search for another video in your native language to respeak into English.

This time, pay extra attention to translation errors.

- Which sub-types of translation error did you make?
- Why do you think they occurred?

# The Future of Intra and Interlingual Respeaking

## Technology

- Training for speech recognition is becoming shorter and Dragon is reaching closer to the desired 100% accuracy.
- Dictionaries of SR programmes expected to be more thorough.
- For software to overtake respeakers speaker-independent SR software is needed.
- Punctuation issues.

# Automatic vs. manual subtitles

Analysis of a sample of clips of TV programmes from a broadcaster

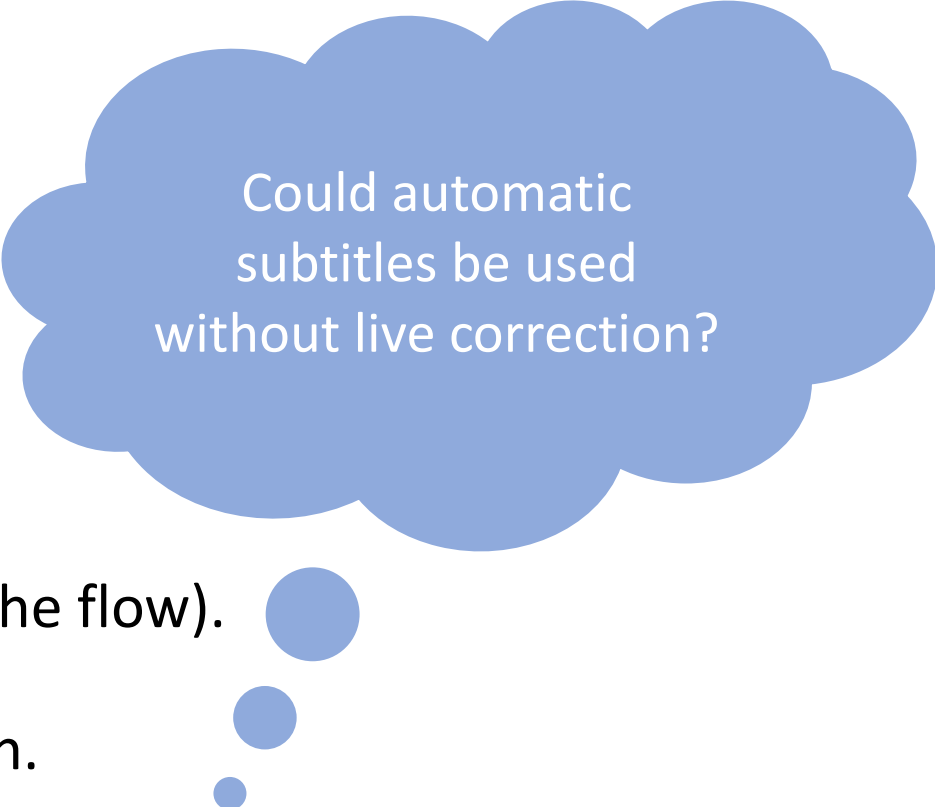
	<b>Respeaking</b>	<b>MS</b>	<b>Google</b>
News 1	98.76%	95.49%	94.08%
News 1 without punct. errors	98.76%	97%	97.89%
News 2	97.37%	97.04%	95.95%
News 2 without punct. errors	97.37%	98%	98.26%
News 3	97.05%	96.08%	94.39%
News 3 without punct. errors	97.05%	97.56%	97.74%

# Automatic vs. manual subtitles

News 4	98.28%	95.86%	95.08%
News 4 without punct. errors	98.28%	97.31%	97.77%
Sports 1	93.87%	95.51%	93.5%
Sports 1 without punct. errors	93.87%	96.34%	96.45%
<b>Average</b>	97.07%	96%	94.6%
<b>Average without punct. errors</b>	97.07%	97.24%	97.62%

# Automatic vs. manual subtitles

- Quality of respoken subtitles is lower than the average broadcast on UK TV, but higher than automatic subtitles produced by MS and Google.
- Quality of some automatic subtitles is impressive:
  - Accurate recognition of structured narration
  - Long stretches of spontaneous speech
- Too many minor errors in automatic subtitles (interrupt the flow).
- Too many standard errors in fast and spontaneous speech.

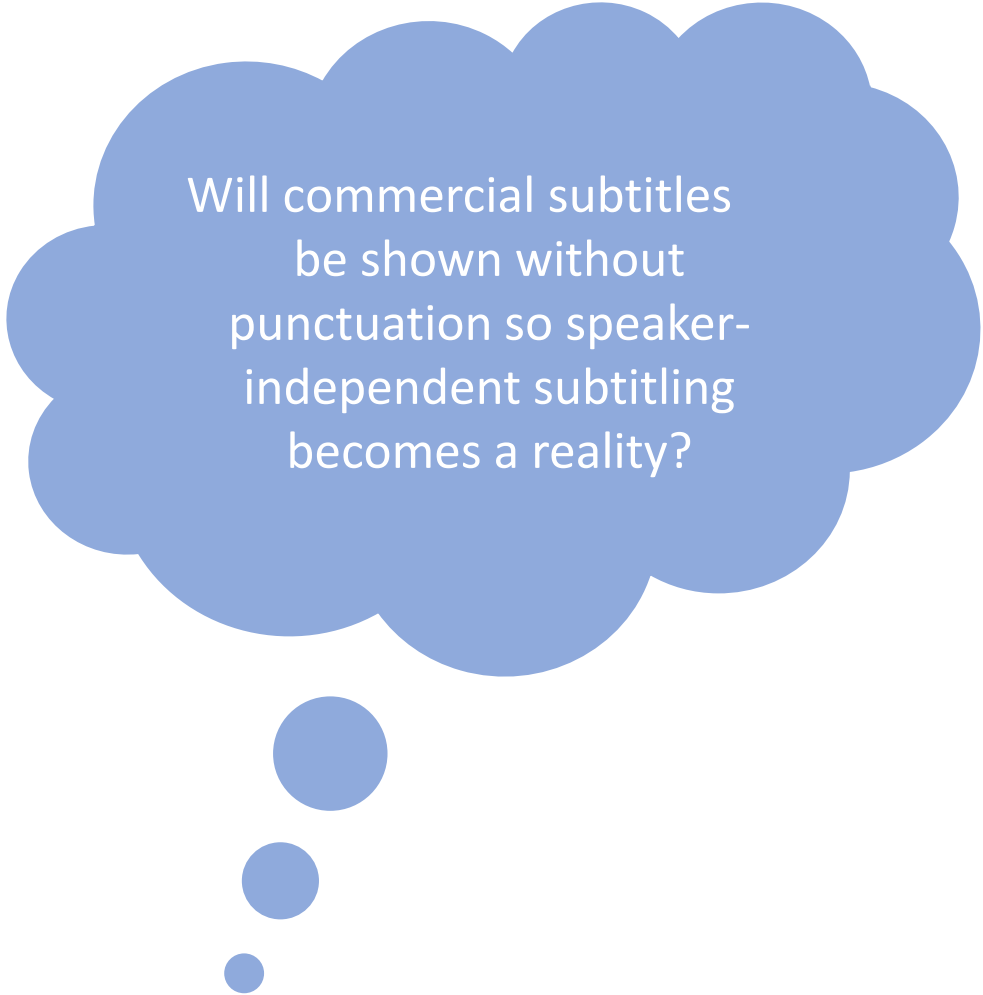


Could automatic subtitles be used without live correction?



# Automatic vs. manual subtitles

- If punctuation is not counted, automatic subtitles are close to the accuracy threshold (some do reach 98%).
- It is unknown if viewers would be happy to put up the minor punctuation errors.
- Software may need to have an automatic punctuation function applied to structured speeches with few speakers.



Will commercial subtitles  
be shown without  
punctuation so speaker-  
independent subtitling  
becomes a reality?

# Current state of interlingual respeaking (professional practice – working conditions)

# Current state of interlingual respeaking (professional practice – working conditions)

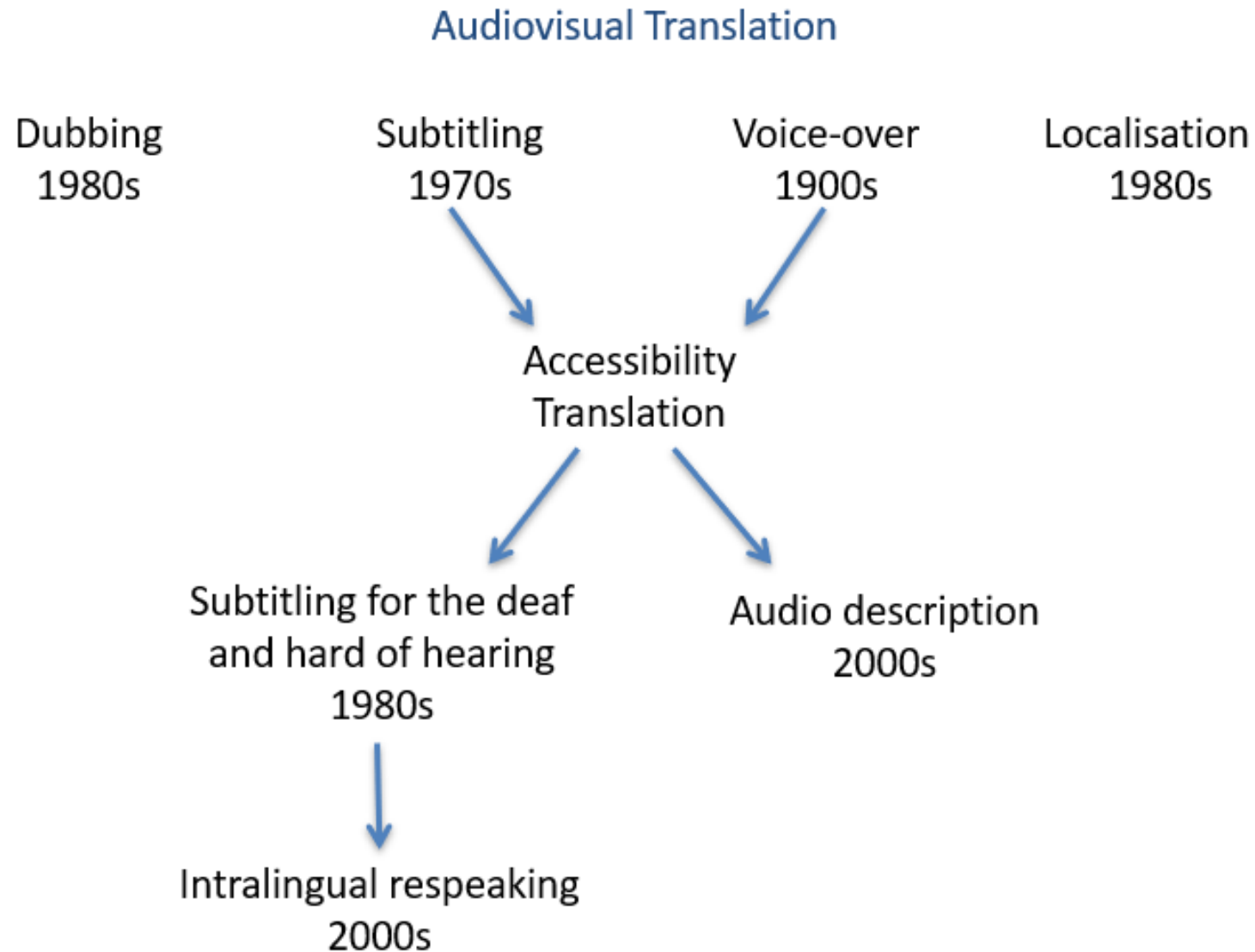
<u>Modality</u>	<u>Recommended rates per country</u>					
	<u>Spain</u>	<u>Italy</u>	<u>Austria</u>	<u>Netherl.</u>	<u>Belgium</u>	<u>Poland</u>
<b>SI</b>	350- 400/half 500/day	300/half 500/day				190- 280/day
<b>Intra</b>	350- 400/half 500/day	300/half 500/day		700/day	350 half 550 full day	234/day
<b>Inter</b>	450/half 700/day	350-450/half 500-750/day	550-600/half 800-1000/day		550–700 day	280/day

# Current state of interlingual respeaking (professional practice – working conditions)

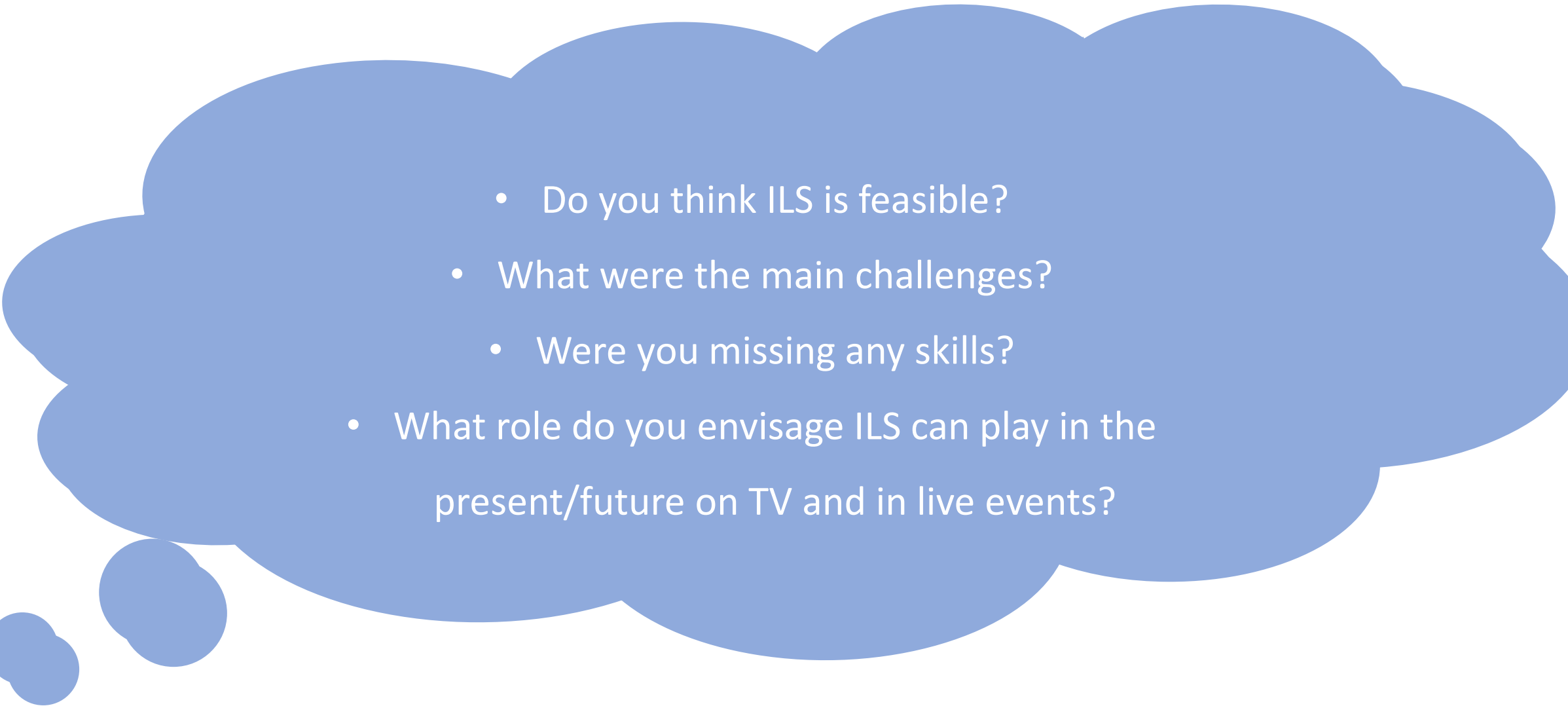
## Other considerations:

- Don't charge per hour, but rather a minimum (i.e. half a day)
- Exclude travel costs
- Intralingual = 80%-100%\* of SI (\*having intra as 100% of SI is possible but in some cases it may lead to employers opting for automatic subtitling).
- Interlingual = 110%-120% of SI
- Price per person/per respeakers, excluding taxes
- Two respeakers for events longer than 40m
- One respeaker per language

# Where does intralingual respeaking sit?



# Closing Discussion

- 
- Do you think ILS is feasible?
  - What were the main challenges?
  - Were you missing any skills?
  - What role do you envisage ILS can play in the present/future on TV and in live events?

# Closing Discussion

- Please keep in touch and help us keep track of any progress being made.

# Closing Discussion

- 
- Thank you!