

Interlingual Respeaking: a comparative analysis of interpreters' and subtitlers' performance

Aims

Overall aim for ILSA: To design, develop, test and validate the first training course for ILS and provide a protocol for this discipline for TV, the classroom and parliament.

Aim of the main experiment:

To train and test participants in an ILS course

To answer the following questions:

- Is ILS feasible
- Who is better suited?
- What are the main challenges?

The experiment: a short online course

A **short online course** was taught to 50 students with backgrounds on subtitling and interpreting covering dictation practice, intralingual and interlingual respeaking.

- **50 students** produced **300 respoken texts** and **100 tests**.
- Texts have been analysed with the NTR model (Romero-Fresco & Pöchhacker, 2017)

Course outline

Week 1 | Dictation Practice

Complete initial long questionnaire, install Dragon, create user profile, example videos of respeaking, intralingual dictation practice

Week 2 | Intralingual Respeaking

Readings on punctuation, condensation and rhythm and intralingual respeaking practice

Week 3 | Interlingual Respeaking

Readings on effort in interpreting, split attention and interlingual respeaking practice

Week 4 | Interlingual Respeaking Test


Interlingual respeaking test and complete final long questionnaire

Course outline cont.

Class Drive folder


Classroom calendar

Google Calendar



Hayley Dawson
Teacher

hayleydawson11@gmail...




Pablo Romero-Fresco
Teacher

pablo.romero.fresco@g...


EN > ES Interlingual Respeaking Course

Add class materials...


Welcome



Instructions.docx
Word




Course Description.docx
Word




Dragon_Commands_ES.pdf
PDF

Week 1 | Dictation




Dictation Practice.docx
Word




Dictation_La accesibilidad a los medios de comunicación audiovisual a través del subtítulo y d...
PDF


Week 3 | Interlingual Respeaking




READING_Split attention- dealing with simultaneous but non-overlapping inputs (Romero-Fresco, ...
PDF




READING_The Effort Models of Interpreting (Gile, 2009).pdf
PDF



Week 3_News_Virtual Reality.mp4
Video



Week 3_Sport_Football match.mp4
Video



Quality Assessment in Interlingual Live Subtitling- The NTR Model.pdf
PDF

Figure 1: Example of Google Class platform

Video clips

Test 1

Speech

Emma Watson gives a speech to the UN about feminism

110 wpm

Test 2

Talking Head

The benefits of gardening on physical and mental health

159 wpm

Main experiment test clip

The screenshot displays a web browser window with a video player and a Word document. The video player shows a woman speaking, with a play button and a progress bar at 0:00. Below the video, there is a 'Speech Details' section with the following information:

- Speech number: 26593
- Duration: 07:25
- Language: (en) English
- Level: Beginner
- Use: simultaneous
- Type: pedagogical material
- Domains: General, Health
- Terminology: proven benefits, stress-relieving, cortisol, sense of purpose/achievement, to thrive, lawnmower, to weed, to prune, muscle groups, osteoporosis, dirt, immune system, skin complaints, psoriasis

To the right of the video player, a Word document titled 'Documento1 - Word (Error de activación de productos)' is open. It contains a table with two columns: 'Skin complaints' and 'Enfermedades cutáneas'. The word 'psoriasis' is written under 'Enfermedades cutáneas'. Below the table, the text 'Diga eso otra vez' is highlighted. The Word document also shows a notification about updates and a status bar at the bottom indicating 'Página 1 de 2', '45 palabras', and 'Español (España)'.

Figure 2: Example of a test clip

The NTR model

(Romero-Fresco & Pöchhacker, 2017)

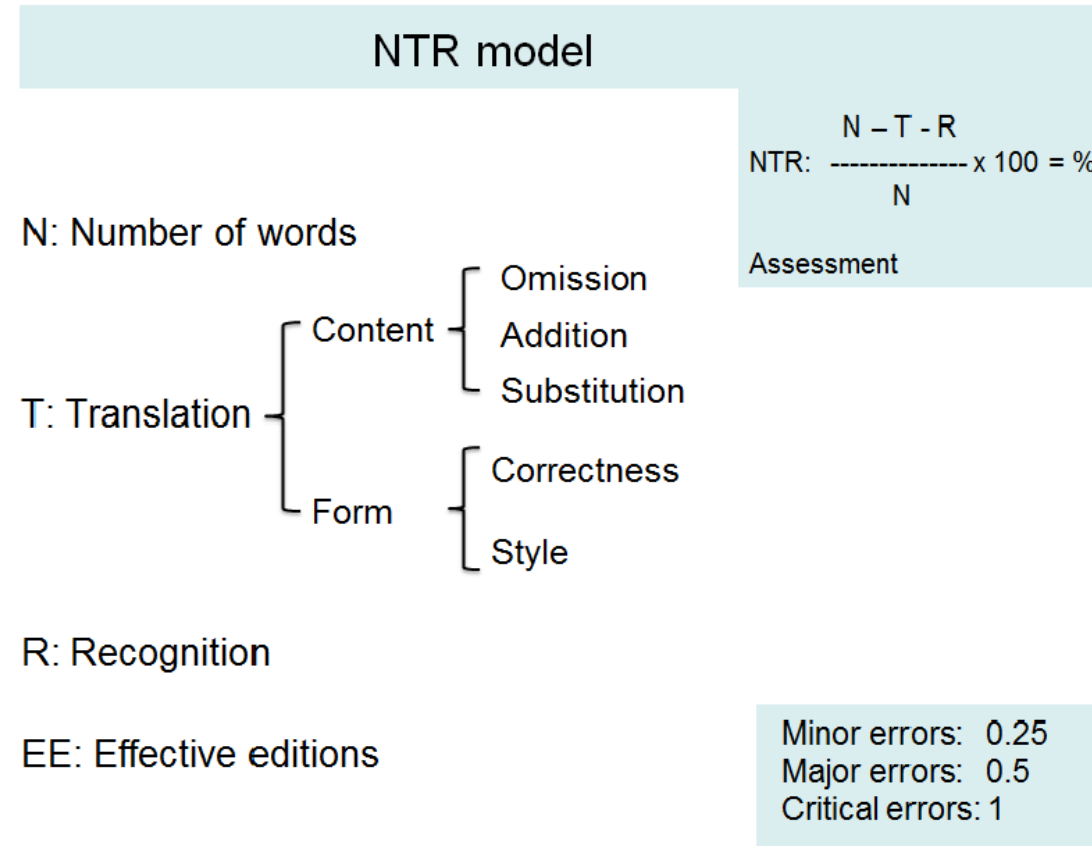


Figure 3: The NTR model (Romero-Fresco & Pöchhacker, 2017)

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 más contenta y (1) mucho más relajada. Por lo tanto, cada vez que me siento algo deprimida lo que hago es salir y pasar algo de tiempo de Servando (desherbando) (2) o plantando flores. Todo esto me hace sentir mucho mejor.	<p>1.MinT (cont-omiss) (0.25): The viewer would miss some of this description.</p> <p>2. MajR (0.5): The error cannot be recognised.</p> <p>EE: No relevant information is lost here.</p> <p>3. MinR (0.25): The error can be recognised.</p> <p>4. MinT (cont-omiss) (0.25): Omission of a dependent idea unit.</p> <p>5. MinR (0.25): The error can be recognised.</p>

Figure 4: Example of NTR analysis for a test

Example of NTR analysis cont.

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>	

Figure 5: Example of NTR assessment for a test

Breakdown of professional profiles

- 22% of participants were subtitlers
- 28% of participants were interpreters
- 46% of participants had experience in subtitling and interpreting
- 4% of participants did not have subtitling or interpreting experience
- 12% of participants had experience in intralingual respeaking

Accuracy rates of tests

	All	Good performers	Poor performers	Interpreters	Subtitlers	Poor interpreters
Accuracy rate	97.6% (4/10)	98.5% (6.5/10)	96.9% (2.5/10)	97.9% (5/10)	97.4 (3.5/10)	97.08% (2/10)

Figure 6: Average accuracy rates

Type of error	All	Good performers	Poor performers	Interpreters	Subtitlers	Poor interpreters
T errors	17.7	13	21.6	16	18.7	19
R errors	17.4	13.4	19.8	16	19	19.2

Figure 7: Average number of translation and recognition errors

Accuracy rates of tests cont.

- Main differences between good and poor performers are errors that impact on full sentences such as omissions and mistranslations and errors that introduce nonsense.
- Good performers were able to contain errors to only affect partial meaning of a sentence.
- Subtitlers incur more omissions, mistranslations and minor recognition errors.
- Poor interpreters have performed worse than the average subtitler.

Type of error	All	Good performers	Poor performers	Interpreters	Subtitlers	Poor interpreters
MinT	13	10.9	14.8	11.9	13.7	13
MajT	3	1.4	4	2.2	3.6	3.2
CritT	1.3	0	2.8	1.7	1.4	2.4
MinR	12.2	10.5	13	10.3	13.9	11.5
MajR	4.8	2.7	6.3	5.1	4.8	7
CritR	0.3	0.2	0.4	0.5	0.2	0.8

Figure 8: Breakdown of severity of translation and recognition errors

Accuracy rates of tests cont.

Type of translation error		All	Good performers	Poor performers	Interpreters	Subtitlers	Poor interpreters
Cont-omiss	Minor	5.5	5.3	8.15	6.6	6.8	7.25
	Major	1.5	0.5	2.1	1.4	1.8	1.9
	Critical	0	0	0	0	0	0
Total cont-omiss		2.3	2.2	3.4	2.6	2.9	3.05
Cont-subs	Minor	2.5	1.8	2.9	1.9	3.3	2
	Major	1.3	0.6	1.7	0.7	1.7	1.1
	Critical	1.6	0.6	2.6	1.5	1.4	2
Total cont-subs		1.8	0.7	2.4	1.4	2.1	1.7
Cont-add	Minor	0.6	0.6	0.6	0.6	0.5	0.7
	Major	0.1	0.1	0.1	0.1	0	0.05
	Critical	0.09	0	0	0.1	0	0.2
Total cont-add		0.2	0.2	0.2	0.3	0.2	0.3
Form-corr	Minor	2.6	2.9	2.7	2.4	2.6	2.25
	Major	0.03	0.05	0	0	0	0
	Critical	0	0	0	0	0	0
Total form-corr		0.9	1	0.9	0.8	0.9	0.75
Form-style	Minor	0.5	0.4	0.4	0.3	0.4	0.4
	Major	0.02	0	0.05	0	0	0.05
	Critical	0	0	0	0	0	0.2
Total form-style		0.2	0.1	0.15	0.1	0.1	0.45

Good performers have:

- 4 times fewer omissions of full sentences
- 3 times fewer mistranslations of full sentences
- 4 times fewer mistranslations that cause change of meaning

Figure 9: Breakdown of the type of translation error

Previous intralingual respeaking experience

- Five participants completed both tests and are poor performers with an average accuracy rate of 97.40%.
- Poor performers with respeaking experience performed better than the other poor performers.
- The two participants with substantial respeaking experience had considerably less R errors than T errors. This suggests software skills are necessary, although T and R errors are equally as important.

Profile	Previous training	Average accuracy rate	T errors	R errors
Subtitler	Self-taught	97.20%	37	16
Subtitler	1 year	97.90%	26	6
Subtitler & Interpreter	1 week	97.17%	18	27
Subtitler & Interpreter	1 month	97.38%	26	18
Subtitler & Interpreter	Less than 1 month	98.52%* 1 text	10	16
Subtitler & Interpreter	1 month	97.38%	24	20

Figure 10: Averages of participants with experience in intralingual respeaking

Subtitlers and interpreters

- Those with previous experience in subtitling and interpreting performed poorer than the average interpreter (97.9%) but better than the average subtitler (97.4%).
- The average length of interpreting training is a 3 – 6 month module at university.

	All	Good performers	Poor performers
Accuracy rate	97.49% (3/10)	98.32% (5/10)	97.16% (2/10)

Figure 11: Average accuracy rates of those with both subtitling and interpreting experience

Type of error	All	Good performers	Poor performers
T errors	17.7	10.4	20.5
R errors	17.2	15.9	17.6

Figure 12: Translation and recognition errors of those with both subtitling and interpreting experience

Bilinguals

- Two participants did not have any previous experience of subtitling, interpreting or respeaking.
- Their average of T and R errors is lower than the average for interpreters and subtitlers.
- More research is required to determine whether bilinguals can be trained in interlingual respeaking.

Emma Watson Speech	Gardening Talking Head	Average	T errors	R errors
98.31%	97.03%	97.67%	15.5	17.5
98.12%	98.35%	98.23%	17	15.5

Figure 13: Accuracy rates and errors of those with no previous subtitling or interpreting experience

Course engagement

- Participants who engaged in the course performed better (just about) than those who did not engage and only submitted work.

	All	Good performers	Poor performers	Interpreters	Subtitlers
Accuracy rate	97.54% (3/10)	98.41% (6/10)	97.41% (3/10)	98.11% (5/10)	97.08% (2/10)

Figure 14: Accuracy rates of those who engaged in the course

	All	Good performers	Poor performers	Interpreters	Subtitlers
Accuracy rate	97.51% (3/10)	98.52% (6/10)	96.94% (2/10)	97.43% (3/10)	97.60% (4/10)

Figure 15: Accuracy rates of those who did not engage in the course

Breakdown of test performance

- Overall 41% of tests reached the threshold of 98%.
- In the Emma Watson speech 19/46 participants (41%) reached the threshold of 98%.
- In the Gardening talking head 18/45 participants (40%) reached the threshold of 98%.

Text	Interpreters	Subtitlers	Subtitlers/ interpreters	Bilinguals	Overall Average
Emma Watson Speech	97.65% (4/10)	97.22% (3/10)	97.46% (3/10)	98.22% (5/10)	97.49% (3/10)
Gardening Talking Head	98.05% (5/10)	97.61% (4/10)	97.48% (3/10)	97.69% (4/10)	97.67% (4/10)

Figure 16: Average accuracy rates of group performance by test

Conclusions

- ILS seems feasible (97.6%, 4/10)
- Interpreters perform better than subtitlers
- Interpreter \neq good performer / Subtitler \neq poor performer
- Translation and recognition are equally important and challenging
- Good performers have around 50% fewer translation and recognition 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.
- Subtitlers seem to struggle trying to keep up with the text, as a result they have more omissions, more mistranslations and more recognition errors.

Is interlingual respeaking feasible?

41% of tests reached the threshold of 98%

The answer is 'yes' providing the following:

- Elaborate training in the form of a 12 week module or course
- Training must incorporate extra dictation and software practice (to master live corrections), plus extra intralingual practice (to master listening and speaking simultaneously)
- Interlingual respeaking practice must focus on both practical and theoretical elements:
 - Practical is to include respeaking practice with different genres of television
 - Theoretical is to include learning and applying subtitling and interpreting theory plus learning how to assess respoken texts



UniversidadeVigo



Thank you for your attention

