The Interpreters' Newsletter 2023 (28), 37-67
DOI: 10.13137/2421-714X/35550 ISSN 1591-4127 (print) ISSN 2421-714X (online)
https://www.openstarts.units.it/dspace/handle/10077/2119

Same interpretation – different gender. Do implicit theories determine the perception of simultaneous interpreting quality?

Esperanza Macarena Pradas Macías
University of Granada, Department of Translation and Interpreting

Emilio Sánchez Santa-Bárbara
University of Granada, Department of Social Psychology

Abstract

This empirical study analyses the impact of a psychological factor, namely implicit theories (IT), on the perception of the quality of a simultaneous interpretation (SI). The study aims to contribute to research into interpreting quality by analysing the interrelationship between recipients' IT and their evaluation of a SI. Research on IT has focused on the answer to questions such as what makes a success or failure in different domains. Building on a perspective initiated in 2019 in which the study of IT was adapted for interpreting quality research, this study presents a new audio manipulation method that enabled the researchers to change the female voice of a German-Spanish SI (Audio A) into a male voice (Audio B), in order to study the impact of gender on quality assessment in SI. The quality expectations of 63 users (laypersons) and 54 interpreter trainees (semi-experts) are also analysed. The results lend empirical support for the IT hypothesis regarding gender-related professionalism. In addition, in accordance with the IT hypothesis, being a layperson or semi-expert moderates the effect of the evaluation of different aspects measured in both audios on the perceived overall quality of the interpreter's work.

Keywords

Interpreting quality, implicit theories, gender, voice, expectations, evaluation, audio manipulation, moderation analysis.

Empirical research on quality in SI was initiated by Bühler in 1986 in answer to the quest for determining quality in interpreting (*e.g.* Shlesinger *et al*. 1997). Her methodology consisted of a questionnaire distributed to IACI (International Association of Conference Interpreters) members to determine their expectations regarding interpreting quality by responding to content and form-related quality criteria (*e.g.* terminology use versus intonation). This article presents an empirical study drawing on implicit theories (IT) from psychology considering this classic duality of content and form and also applying the two-pronged approach to analysing quality in interpreting. Initiated by Collados Aís (1998), this kind of analysis gave rise to what Pöchhacker (2013: 39) called "the Granada paradigm in experimental research on interpreting quality". In his words (*Ibid*), this survey-based approach "links up interpreting service users' generic expectations of interpreting quality with their judgments of a given interpreting performance," thus allowing a very clear "distinction between expectations and assessment".

This double perspective enables researchers to demonstrate that initial quality expectations of users and experts do not necessarily coincide with their evaluation of a given interpretation. What is more, approaching the analysis of interpreting quality evaluation by comparing the evaluation of a given interpretation by service users/clients and/or experts has revealed the important role that non-verbal factors play in the evaluation of interpretations (see Collados Aís *et al*. 2007).

Non-verbal factors may be very wide-ranging and thus also encompass psychological aspects. In this sense, theories in psychology can help us to gain further insight on quality evaluation. Precisely for this reason, the authors applied IT to interpreting quality research (see section 1.3.). The study of IT arose in relation to personality, intelligence and morality (Dweck *et al*. 1993). For decades, IT have sought to find answers to questions such as what leads to success or failure in different fields. IT impact on the goals, motivation, attributions, expectations and behaviour of participants in a certain context, and affect both the self-perception of a person and his or her perception of others.

With this in mind, and in order to explore the interrelationship between recipients' IT and their evaluation of a SI, the authors designed the present correlational study (see section 2).

The analysis considers the quality expectations of 63 laypersons (potential interpreting service users) and 54 interpreter trainees, as well as their evaluation of one of 2 SI (Audio A, original SI rendered by a woman or Audio B, a manipulated version of Audio A so as to seem to be rendered by a man). These data, obtained by means of a survey carried out in the University of Granada (Spain) (see section 3), are discussed in the final sections, with a view to ascertain why laypersons and semi-experts evaluate in the way they do and, last but not least, to identify the IT of potential interpreting service users and interpreter trainees with respect to the profession itself and to the profile of the interpreter.

1. Interpreting quality research

As mentioned in the introduction, Bühler (1986) launched the first quality survey on expectations in conference interpreting. In her questionnaire, she asked members of AIIC to rate the importance of various linguistic and extra-linguistic (pragmatic) criteria for an interpretation. And, as Zwischenberger (2010: 128) puts it, "her list of criteria became something like the backbone of empirical research into quality". The main aim of the present study is to determine which factors predict the quality of SI and which characteristics define the interpreter. Accordingly, we will focus on a review of those criteria that are indicators of quality, according to the many descriptive analyses of empirical studies (for a review, see *e.g*. Collados Aís *et al*. 2011).

A consistent finding in survey studies has been that respondents consider the quality of an interpretation to lie above all in the correctness of content criteria such as the following: *correct and complete transmission of the information of the original speech* (also called *faithfulness to the original speech*), *terminology* and *logical cohesion* (*e.g*. Collados Aís *et al*. 2011). At the same time, other criteria such as *voice*, *intonation* or *accent* are considered less important. Following the sociological definition of "expectations" as a preconceived idea based on personal beliefs about how an individual or group should behave and what is likely to take place in the future, it is understandable that laypersons as well as experts associate the most complete and correct reproduction possible with the quality of the interpreted discourse. Non-verbal or formal criteria have always been regarded as least important. Specifically, *voice* and *accent* tend to be relegated to the last positions, although with some differences between the groups of laypersons (Kurz *e.g*. 1993) and interpreting experts (Bühler 1986). The latter have attached more importance to voice than laypersons. Collados Aís (1998) was the first researcher to explore whether the expectations that an interpreting service receiver expressed *a priori* coincided with the evaluation s/he made after listening to a given interpretation. In her study, respondents (interpreters and users) were asked to evaluate one interpretation with errors of content but with a lively intonation, and another with no errors of content but with a monotonous intonation. The results led her to conclude that a formal aspect, i.e. a non-verbal criterion such as *intonation*, acquired far greater importance in the actual evaluation than the interpreters and users had indicated in their expectations.

Shortly afterwards, in 2000, Collados Aís formed the ECIS research group. In the course of diverse projects, ECIS followed the same dual methodology to investigate other formal criteria commonly included in quality research: *accent*, *diction*, *fluency*, *style* and *voice*. All these studies (Collados Aís *et al*. 2007) confirmed that the priority ranking of content and formal criteria changes when users actually evaluate a specific interpretation.

In this respect, the criterion *fluency* is particularly striking. In expectations, respondents almost always place it in the middle of the priority scale (cf. Collados Aís *et al*. 2007), yet its impact on evaluation is important (*e.g*. Pradas Macías 2003 or Rennert 2019). Pradas Macías (2003) conducted an extensive interdisciplinary review of the concept of fluency without finding a definition of its nature, while Chiaro and Nocella (2004) even stated that it is practically impossible to associate fluency clearly with either category, *i*.e. that of verbal or non-verbal criteria.

As indicated by Pradas Macías (2019), this fact suggests that fluency has a dual nature: while it is perceived through the occurrence of what may be considered representative elements of fluency, such as pauses, hesitations, *etc*. (*e.g*. Pradas Macías 2003), these elements seem at the same time to have a psychological effect on the listener, influencing his or her perception of the content.

In the light of such findings, researchers have increasingly turned to psychology to unravel what influences the evaluation of quality in SI.

A case in point is García Becerra (2012), who carried out a study in which diverse groups of respondents were asked to assess the same series of interpretations (from high to low quality) but presented in a different order (*Ibid.*). She drew specifically on the theory of the *priming effect* (when an individual's exposure to a certain stimulus subconsciously influences his or her response to a subsequent stimulus). The results showed that receivers alter their evaluation of a particular interpretation depending on which interpretation they have heard before.

1.1 The influence of gender on interpreting quality

As indicated by Du (2020), compared with the amount of work on gender in Translation Studies, very little research has been conducted to explore gender-related issues in interpreting. Moreover, this research perspective is relatively recent, arising in the 21st century.

Nevertheless, various approaches to the topic of gender in Interpreting Studies have been undertaken. Most of these have centred on gender-based *differences* in interpreting (*e.g*. Magnifico/Defrancq 2015; Russo 2016; Cámara Aguilera/Pradas Macías 2019; Xiaoshu *et al*. 2021), but there are also studies which take a historiographical perspective (*e.g*. Baer/Massardier-Kenney 2015; Defrancq *et al*. 2021). Other scholars have centred on quality and the associated expectations regarding the interpreter's role (Zwischenberger 2013) and, last but not least, on the incidence of gender on interpretation recipients (*e.g*. López Varela 2016; Christen y Gracia 2020).

We will briefly dwell on the research by Christen y Gracia (2020) because it deals in a similar way to this article with the impact of a male or female voice on the quality of a performance. Her focus lies specifically on the need for interpreters to be aware of the importance of vocal hygiene for health, timbre and vocal range. 63 subjects participated in her study evaluating the recorded performances of 4 interpreters: 2 novice interpreters (male and female) and 2 experienced interpreters (male and female). The sample of listeners was stratified by age group and gender. As Christen y Gracia (2020) points out, "the findings reveal the respondents' preference for male and novice interpreters. Female voices were judged more harshly, especially those of older interpreters".

1.2 IT in psychology and their application to interpreting quality research

The study of IT arose in relation to personality, intelligence and morality (Dweck *et al*. 1993). Encompassing elements of social psychology research such as personali-

ty attributes, IT refer to the ideas of collective groups regarding how a professional should act within his or her sphere of activity.

The IT approach has had and continues to have a great impact in different areas of psychology – not only in social psychology – from the initial publications, *e.g*. Dweck and Elliot (1983), to their extension of the model (Dweck *et al*. 1995) to include other facets such as stereotypes, cross-cultural differences and the origins of IT.

The origins of the approach go back to Kelly's (1955) studies on personal constructs or theories of non-experts – the ones we use in our daily lives – and how these affect perceptions of self and others. The approach has been applied to many research areas, including personality, intelligence, interpersonal and affective relationships, emotion, leadership, sports performance, creativity, moral character or ability (for a review, see Dweck 2012).

IT are basic assumptions that people have about themselves and the world. These beliefs guide a person's goals, either towards evaluating, judging and measuring an attribute when it is considered fixed, or towards developing the attribute or understanding its dynamics when it is considered malleable.

IT have also been studied with promising results in the field of translation and interpreting, as well as in language learning (Ryan/Mercer 2012; Enríquez Raído 2018; Chodkiewicz 2020). Research also provides insights into the impact of gender on IT, resulting in no (or small) gender differences in abstract conceptions, but larger differences between men and women when people perceive the phenomenon in real-life contexts (Glück *et al*. 2009).

Sánchez Santa-Bárbara and Pradas Macías (2019) included IT and theory about stereotypes in a study aiming to link up users' generic expectations with their judgments of a given interpreting performance. The aim was to establish if there is a difference between the beliefs and expectations of users with more or less experience of interpreting services. Following the logic of Roach (1978), users without experience are likely to categorise interpreters on the basis of exemplars (*e.g*. a specific interpreter), while more experienced users will resort to abstractions (more *elaborated* mental syntheses). Such processes of abstract categorization are characteristic of experts in a specific field of knowledge. This in turn led to the hypothesis that the IT of professional interpreters and users are likely to be different, as is also the case with trainee interpreters and potential users of professional services; and that these differences will depend on their degree of exposure to the field of knowledge.

In addition, the researchers added a new element to the theoretical framework, namely the gender stereotypes by which certain psychological and behavioural characteristics are assigned differently to men and women.

The most surprising result of this study was that the participants evaluated the interpretation differently depending on whether the interpreter was a man or woman, with a more favourable evaluation for the man. This finding encouraged us to explore the influence of the interpreters' gender in greater detail in the present study.

2. Methodology

2.1 Research question and hypotheses

IT enable subjects to identify characteristics associated with a certain domain (in this case, interpreters) and to develop expectations about how these professionals will act and behave. These IT are acquired through experience in a particular setting.

According to this approach, laypeople and practitioners may have different IT about interpreters' skills, because they have had different knowledge and expertise in relation to this profession. In line with this premise, the hypotheses are formulated as follows:

H1: To the extent that the interpreter's verbal and non-verbal behaviour matches the service users' expectations as well as other characteristics contained in their prototype, the interpreter's work will be perceived as highly effective.

H2: IT about the interpreting profession may have an associated gender profile. Depending on this association, the evaluation of an interpretation will be different depending on whether it is performed by a man or a woman.

2.2 Objectives

The descriptive objectives of this study are:

O1: To analyse what semi-experts (interpreter trainees) and laypersons expect from an interpreter and an interpretation, and to correlate their expectations with their perceptions of overall quality.

O2: To determine whether the evaluation of the quality of the interpreter's work differs depending on the gender of the interpreter and if so, which criteria are significantly affected.

O3: To analyse the evaluation rendered by semi-experts and by laypersons in contrast with their appreciation of quality expectations.

2.3 Participants

To achieve these objectives, a correlational cross-sectional study was carried out to identify the content of interpreter expectations and the evaluation of a given interpretation. 117 students of the University of Granada participated in the study, although the final sample is made up of 116, as one student did not answer the question regarding his/her gender. 54 subjects were from the Faculty of Translation and Interpreting and are considered semi-experts in the framework of this study (11 men and 43 women). The rest of the sample consists of 62 students from the Faculty of Political

Sciences and Sociology who are considered laypersons in this study (19 men and 43 women). The participants were recruited expressly on the basis of their differing levels of knowledge of the professional work of an interpreter. The criterion for selecting the group of semi-experts was that the interpreter trainees had already taken the core courses in interpreting of the Bachelor's Degree in Translation and Interpreting.

2.4 Material

For evaluation by the 2 groups, 2 audio recordings of a SI were prepared. The material was selected from the ECIS corpus (ECIS Research Group 2003) and consisted of the interpretation into Spanish of a speech delivered by a female German politician in a session of the European Parliament. The interpretation (Audio A), performed by a female interpreter, had 1,185 words, a duration of 568 seconds and a speed of 125.18 w/m. It was chosen because in a previous evaluation by experts it had achieved 4 points for the overall quality of the interpretation on a five-point Likert scale, this being the most positive evaluation by experts for a female interpretation in the corpus[1].

To achieve the greatest correspondence between the 2 recordings, the female voice was modified so that it sounded like that of a man (Audio B). This was done by using Audacity, the free audio editing program, and manipulating the option *Effect,* specifically *Change tone*[2], selecting a reduction of -19% in the pop-up window[3].

To ensure that recipients do indeed perceive Audio B as rendered by a man, the authors asked potential interpreting service users from their faculties to indicate whether they would use the female or male interpretation for research purposes. To do this, they played different segments of Audio A and Audio B, so that they would not be recognised as being identical. The potential users expressed no doubt as to which gender was associated with the voice in each version.

Nevertheless, for maximum assurance, after the survey the subjects were asked to indicate by a show of hands who had heard the version with the female voice and who had listened to the male version. The results coincided with the researchers' data on the distribution of the 2 audios.

Subjects voluntarily agreed to participate in the survey[4]. In a first step, the expectations questionnaire was distributed. After completing the questionnaire, participants were given receivers with headphones, one containing the recording of version A and

1 The audios of the ECIS Research Corpus were evaluated by interpreters as part of the research project *Evaluación de la calidad en Interpretación Simultánea: Parámetros de incidencia* funded by the Spanish Ministry of Science and Technology (ref. BFF2002-00579).

2 The main difference between the voice of a male and a female is the Fundamental Frequency (F0) (see García/Tapias Merino 2000).

3 We would like to express our special thanks to the laboratory technician at the Faculty of Translation and Interpreting, Juan Carlos Rodríguez García, for this technological achievement.

4 To test the method beforehand, on the 1st March 2022 a pilot study was carried out with 2 groups of SI students, one with English as their B language and the other with French, enrolled in the Degree programme in Translation and Interpreting at the University Pablo de Olavide (Seville, Spain). The aim of this pilot study was to confirm the suitability of the survey material and procedure.

the other of version B. Each version was randomly distributed to one half of the group. It was felt that subjects should receive the SI as is usual in conference rooms, *i.e*. with receivers and headphones, rather than in a laboratory, which would have added an element of artificiality. The researchers also avoided artificially editing the audio fragment, as this might have been perceived as an attempt to manipulate the audio. Instead, they explained that after 2 or 3 minutes they would give a signal to indicate that the participants should stop listening to the recording. The explanation was that this duration was sufficient to get an idea of the content and presentation of the performance.

After listening to the interpretations (A or B), the researchers distributed the evaluation questionnaire and asked participants to indicate on the first sheet the number of the receiver they had listened to. This way, the researchers could identify the questionnaires that evaluated audio A or audio B. Furthermore, both questionnaires (expectations and interpretations) could be matched with the 5 final numbers of each subject's mobile phone, which was a data required in both.

2.5 Survey on SI expectations and evaluation

As mentioned above, the survey included a questionnaire[5] about expectations and another one to evaluate the interpretation (Audio A or Audio B). The expectations survey was distributed at the start of the study and featured 13 aspects related to the extent to which interpreters are likely to be *x*…; together with 14 further skills and characteristics (see Appendix). The subjects indicated their answers on a seven-point Likert scale.

Subsequently, the participants listened to version A or B and then filled in the evaluation questionnaire. In the first part, they were asked to assess 13 aspects of the interpreter on a seven-point Likert scale ranging from negative to positive (see Table 1). The second set of questions covered 14 quality criteria related to the interpretation (see Table 2). The questionnaire included the following aspects as representative of 7 evaluation criteria regarding the interpreter and the interpretation: *profile*, *voice and appearance*, *form*, *fluency*, *time-lag*, *presentation style* and *perception* of the interpretation. The negative/positive extremes of each item were alternated to avoid the tendency to mark one side or the other automatically.

5 The expectations questionnaire was designed by the group of researchers in the framework of a broader project, QINV: http://qinv.ugr.es/index-en.htm. The list of criteria related to skills, characteristics and inherent qualities is based on the consolidated questionnaire design of ECIS (http://ecis.ugr.es/index-en.htm) and on methods commonly used in social psychology.

I found the interpreter (Profile)	- unconfident/confident - expressive/inexpressive - not very skilled at communicating the ideas effectively/highly skilled at communicating the ideas effectively - highly competent/not very competent - not very reliable/highly reliable - highly professional/not very professional - highly decisive/not very decisive
The interpreter (Voice and appearance)	- had an unpleasant voice/pleasant voice - acted calmly under stress/acted nervously under stress
The presence of the interpreter was (Appearance)	- scarcely noticeable/highly noticeable
The interpreter's intonation was (Form)	-monotonous/lively
The interpreter's articulation was (Form)	- not very clear/very clear
The interpreter's pronunciation was (Form)	- native/non-native

Table 1. Aspects of the profile and professionality of the interpreter

The speed of the interpretation was (Form)	- slow/fast
The rhythm of the interpretation was (Form)	- not very fluent/very fluent - constant/interrupted
In the interpretation there were (Fluency)	- many pauses/few pauses - few hesitations/many hesitations - pauses at inappropriate moments/pauses only at appropriate moments - complete sentences/incomplete sentences
The duration of most of the pauses was (Time-lag)	- short/long
The register was (Presentation)	- highly appropriate to the situation/not very appropriate to the situation
I found the expression in the interpretation (Presentation)	- confusing/clear
I found the interpretation (Perception)	- difficult to understand/easy to understand - not very coherent/highly coherent - difficult to follow/easy to follow - not very faithful to the original speech/ highly faithful to the original speech

Table 2. Quality criteria related to the SI

Following the classification initiated by Bühler (1986), we distinguished between criteria related to the content of the interpretation (*e.g.* expression, clarity, adaptation to the user) and others related to the form, essentially fluency, intonation, voice and accent.

At the end of the questionnaire, participants were asked to mark the overall quality of the interpretation from 1 to 10.

2.6 Analysis procedure

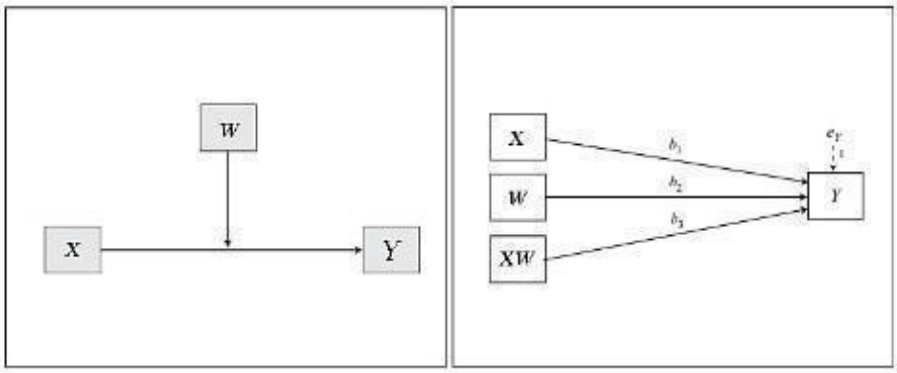
In order to ensure that the 2 groups did not differ in other variables that could affect the objectives, different mean comparison and contingency table analyses were performed.

To predict the value of the quality of interpretation from different explanatory variables, the authors used a procedure which allowed them to work with a reduced set of relevant variables. First, descriptive and bivariate analyses were carried out in order to test for similarities and differences between the 2 groups.

Secondly, a multiple linear regression analysis was carried out to predict the overall quality of the interpretation on the basis of the dimensions or factors taken from the recipients' evaluations. Next, a multiple linear regression analysis was performed using the "stepwise" procedure, so that all variables which did not meet the criterion of statistical significance were excluded.

In addition, an analysis of simple moderation with a moderating qualitative variable was run with the SPSS version of PROCESS (Model 1) to test whether the moderating variable (W = Status of the evaluator: Laypersons versus Semi-experts) moderates the impact of the effect of the independent variable (X = perception of the interpretation listened to) on the dependent variable (Y = perceived quality). In other words, PROCESS was used to carry out a multiple regression analysis to obtain 3 coefficients: $\beta 1$, which quantifies the principal effect of X on Y; $\beta 2$, which quantifies the principal effect of W on Y; and $\beta 3$, which quantifies the principal effect of the interaction between X and W on Y.

To fulfil the requirements of the statistical software PROCESS (see Figure 1), X (Evaluation) was measured as the median of the values given to all characteristics of the interpretation and was then transformed to obtain a measure as a dichotomous variable (with -0.5 codified as lower and 0.5 as higher). W (Status) is naturally dichotomous (Laypersons/Semi-experts) and Y (Overall Quality) was also dichotomized (with -0.5 codified as lower and 0.5 as higher). In order to dichotomize variables X and Y, the median of each variable was used to divide the scores between higher and lower.



Conceptual and statistical diagram for a simple moderation analysis
Source: Hayes (2013)

Figure 1: Statistical software PROCESS

3. Results

3.1 Results for expectations

The means analysis of the expectation responses yields some statistically significant results which allow us to draw some conclusions about differences in expectations according to gender and status group (laypersons and semi-experts).

The results (see Appendix 1) reflect that the semi-experts in general have higher expectations of the interpreters than the laypersons in the following items: [interpreters are likely to] 'be competent', 'be flexible', 'be resolute', 'be audacious', 'go unnoticed', 'adapt to the user', 'show confidence' and [the interpretation is] 'fluent', 'faithful to OS' and 'easy to understand'. The only item in which laypersons have higher expectations than the semi-experts is 'have a monotonous intonation'.

With regard to men and women, the scores are evenly distributed and there are no significant differences.

3.2 Results for evaluation

3.2.1 Evaluation of *profile*

Figure 2 shows the results of participants' evaluation depending on their gender or status. Statistically relevant values show that men (5.61) consider the female interpreter more skilled at communicating ideas effectively than women do (4.82) and that laypersons consider her less competent (3.90) than semi-experts do (2.96).

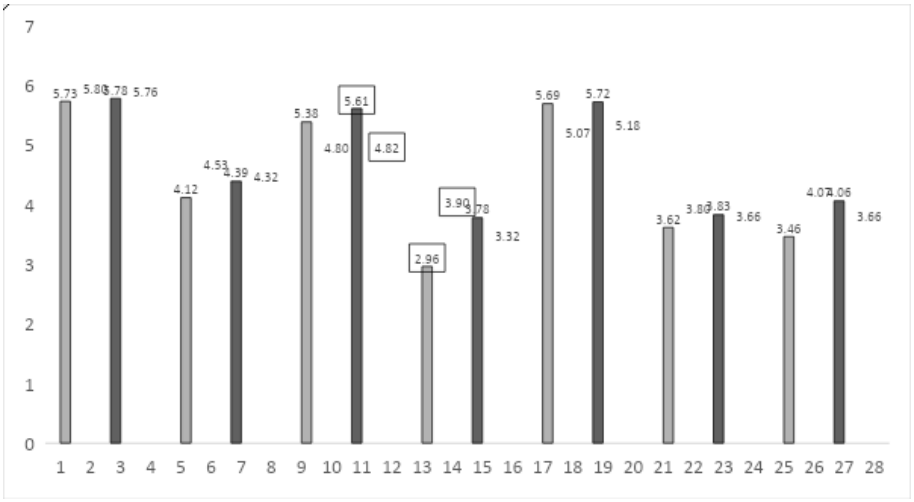


Figure 2: Evaluation of *profile* in the female interpreter's performance

In Figure 3 the most salient result is that men consider the male interpreter less confident (5.75) than women do (4.79) and that laypersons consider the male interpreter less competent (4.39) than semi-experts (3.39), although both status groups perceive the male interpreter as more competent than the female interpreter.

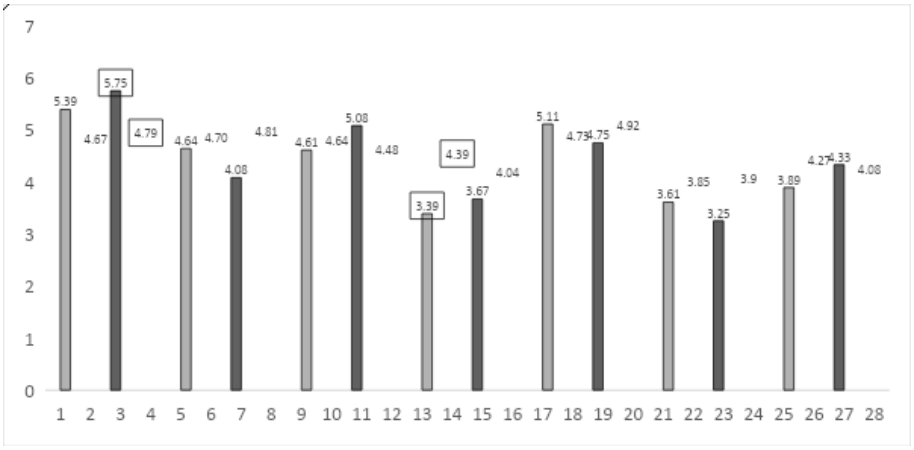


Figure 3: Evaluation of *profile* in the male's performance

3.2.2 Evaluation of *voice and appearance*

The most salient result in Figure 4 is that semi-experts evaluate the voice of the female interpreter as considerably more pleasant (5.65) than laypersons do (4.9).

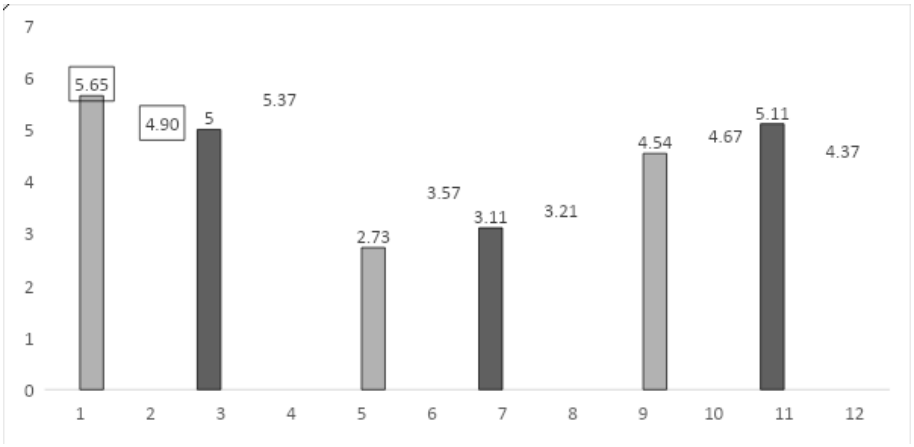


Figure 4: Evaluation of *voice and appearance* in the female interpreter's performance

In the case of the male interpreter, Figure 5 again shows that semi-experts clearly evaluate the voice more positively (5.71) than laypersons do (4.64). Moreover, while semi-experts provide a very similar evaluation of the female voice (5.65; see Figure 4) and of the male voice (5.71; see Figure 5), laypersons evaluate the male voice worse (4.64; see Figure 5) than the female voice (4.90; see Figure 4).

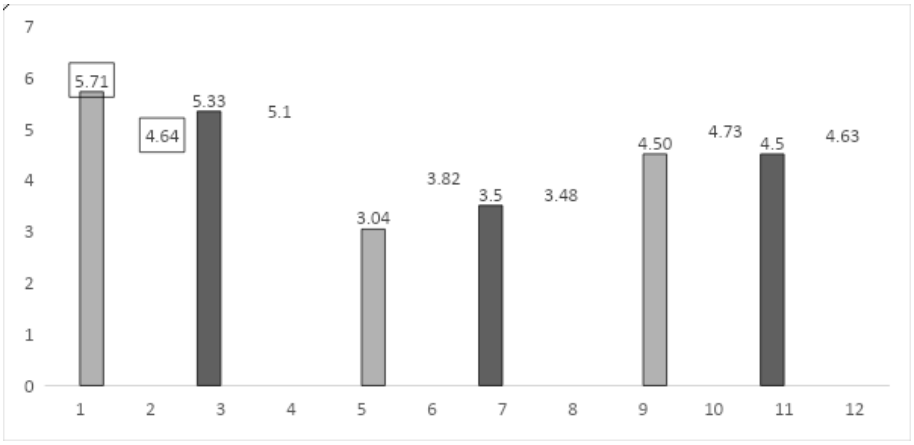


Figure 5: Evaluation of *voice and appearance* in the male interpreter's performance

3.2.3 Evaluation of *form*

The evaluation of non-verbal criteria included in the tag *form* (see Figure 6 and Figure 7) shows that semi-experts attribute nearly the maximum value of 7 to native accent in both interpretations (6.84 and 6.75). In the case of the female interpreter's output, semi-experts also consider the interpretation to be very clear (6.28).

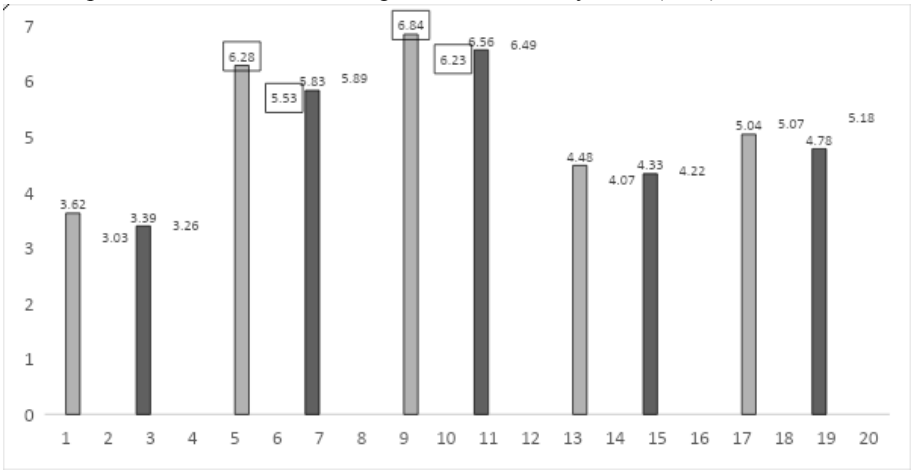


Figure 6: Evaluation of *form* in the female interpreter's performance

In the male interpreter's performance (see Figure 7), semi-experts consider the interpretation significantly more fluent than laypersons do (5.46 versus 4.36).

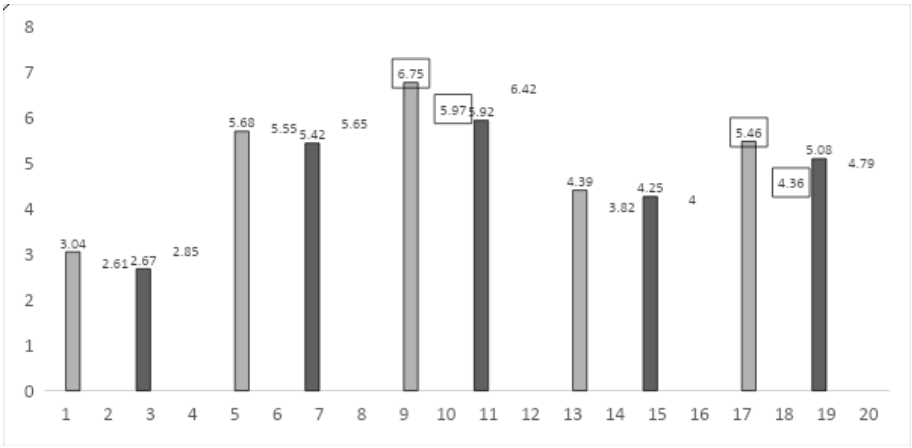


Figure 7: Evaluation of *form* in the male interpreter's performance

3.2.4 Evaluation of *fluency*

In Figures 8 and 9 no results with significant values were reached. Most evaluation values are around 5. The only remarkable aspect is that laypersons consider that there are many hesitations in both performances (values below 4).

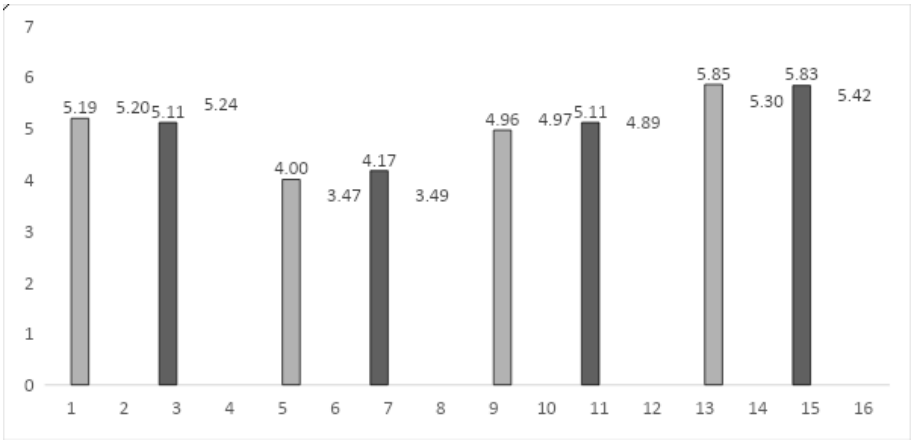


Figure 8: Evaluation of *fluency* in the female interpreter's performance

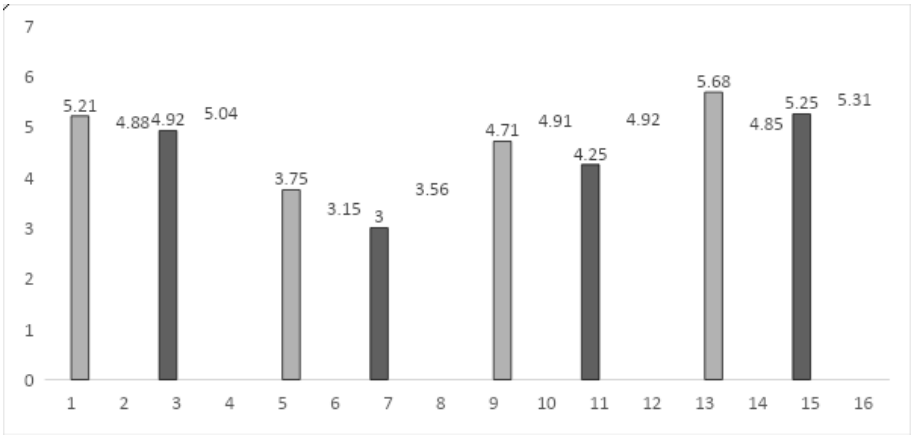


Figure 9: Evaluation of *fluency* in the male interpreter's performance

3.2.5 Evaluation of *time-lag*

Figures 10 and 11 show that there are significant differences in the evaluation by semi-experts and laypersons with regard to *time-lag*. Laypersons indicate exactly the same rating (3.27) for the female and male interpretation, and consider the pauses longer than semi-experts do (2.1 and 2.11).

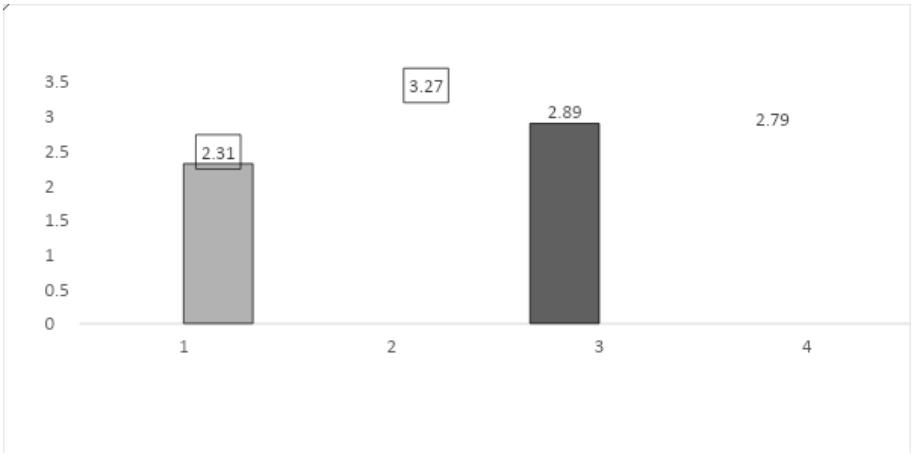


Figure 10: Evaluation of *time-lag* in the female interpreter's performance

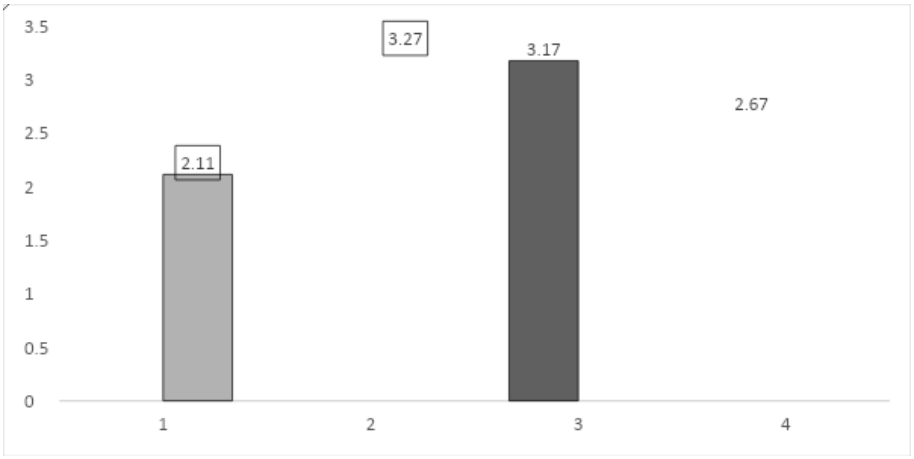


Figure 11: Evaluation of *time-lag* in the male interpreter's performance

3.2.6 Evaluation of *presentation*

Figures 12 and 13 show no significant differences for *presentation*.

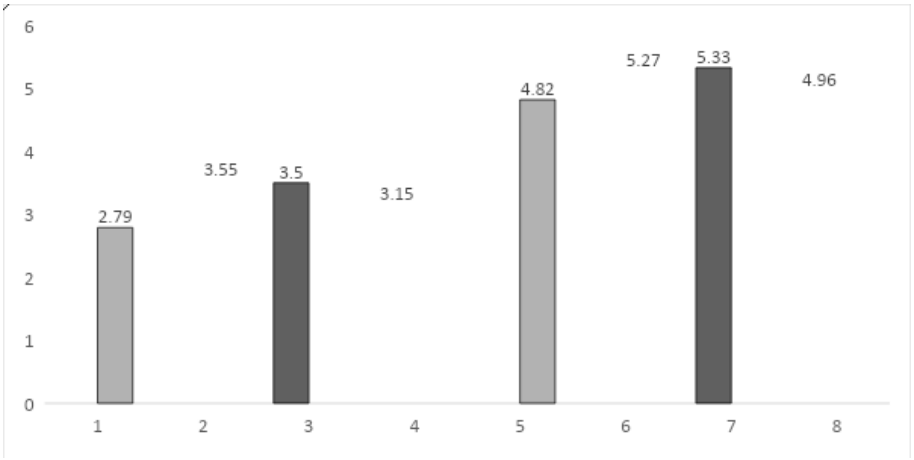


Figure 12: Evaluation of *presentation* in the female interpreter's performance

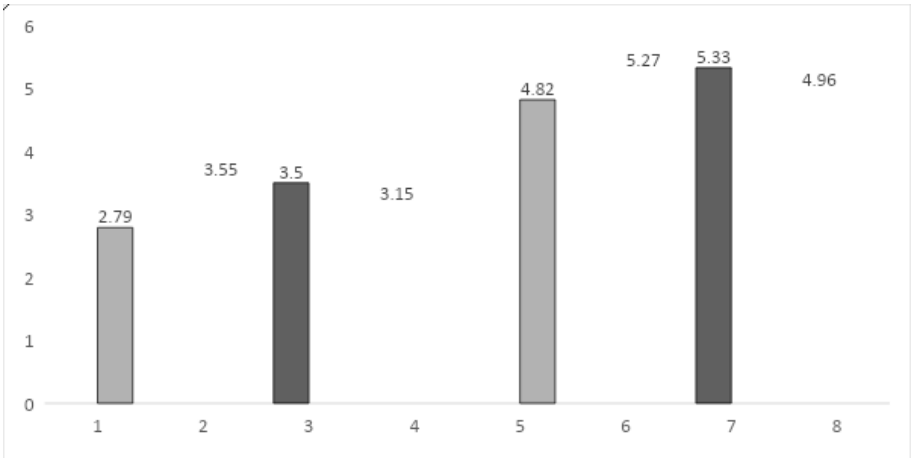


Figure 13: Evaluation of *presentation* in the male interpreter's performance

3.2.7 Evaluation of *perception*

Figure 14 shows a significant result in the evaluation of *perception* of the female interpreter's performance for *overall quality*. While semi-experts rate it 8.09, laypersons rate it 7.07.

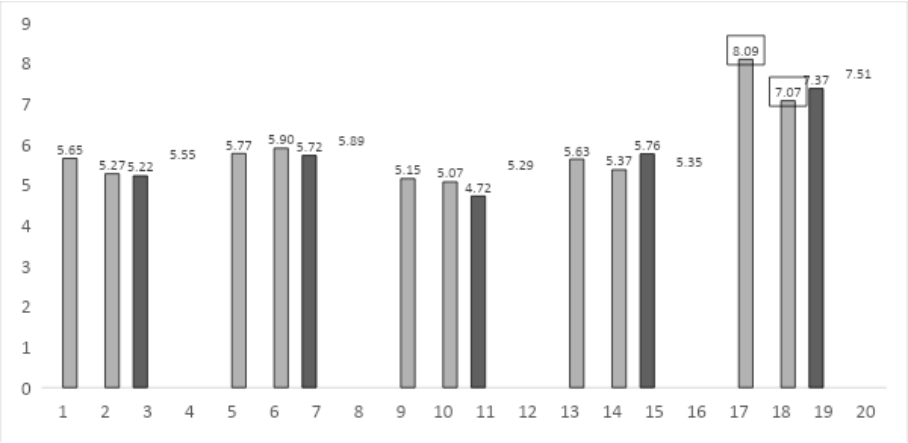


Figure 14: Evaluation of *perception* of the female interpreter's performance

Regarding the evaluation of *perception* of the male interpreter's performance, Figure 15 shows significant values for *difficult to understand / easy to understand*.

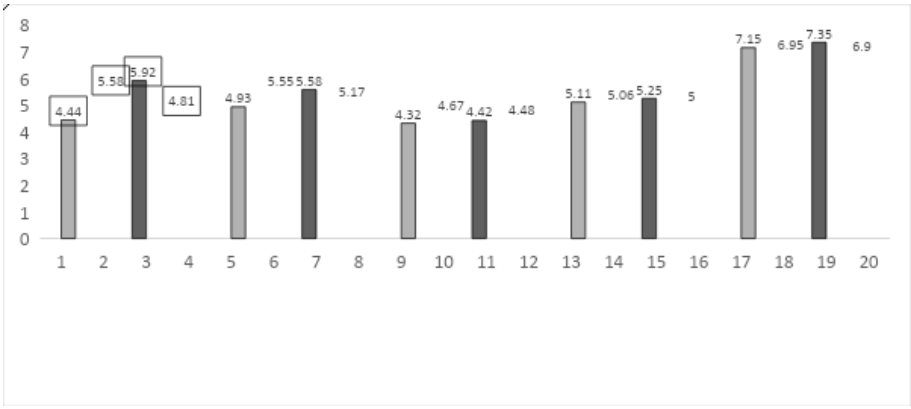


Figure 15: Evaluation of *perception* of the male interpreter's performance

3.3 Expectations correlated with overall quality

In order to operationalize the data for IT, the results from the expectation survey were correlated with the results of overall quality. The analysis was carried out separating the groups (semi-experts and laypersons).

In the case of semi-experts, statistically significant correlations were observed for the aspects 'interpreters are likely to be confident' (r= 0.349, p= 0.023), 'interpreters have specialized knowledge' (r= 0.321, p= 0.023), 'interpreters are likely to go unnoticed' (r= -0.322, p= 0.043) and 'interpreters are likely to speak fast' (r= -0.318, p= 0.049).

Regarding the group of laypersons, the only significant result was obtained for the aspect 'interpreters are likely to have a monotonous intonation' (r= -0.340, p= 0.039).

3.4 Results of the regression analysis

The regression analysis performed to predict overall quality was based on 5 evaluation dimensions: *Form, time-lag, fluency, perception and profile* (see Table 3). The results show that these dimensions predict 68% of the *overall quality* (R2= 0.680; p= 0.000; see Table 4).

Model	R	R2	Adjusted R2	Standard Error
1	0.825a	0.680	0.664	0.9316

Predictors: (Constant), *Form*, *Time-Lag*, *Fluency*, *Perception*, *Profile*

Table 3. Predictive evaluation dimensions

Model				Standardized Coefficients Beta	t Lower Limit (LLCI)	p Upper Limit (ULCI)	Confidence Interval 95,0% for B	
1	(Constant)	1.437	0.857		1.676	0.097	-0.264	3.139
	Fluency	0.285	0.110	-0.163	2.587	0.011	-0.504	-0.066
	Perception	0.620	0.086	0.490	7.209	0.000	0.449	0.791
	Profile	0.326	0.100	0.229	3.268	0.001	0.128	0.524
	Time-Lag	0.125	0.064	0.125	1.967	0.052	0.001	0.251
	Form	0.277	0.139	0.129	1.998	0.049	0.002	0.553

Table 4. Dependent Variable: *Overall quality*

All 5 of these dimensions of *evaluation* are statistically significant for predicting the dependent variable *overall quality*: *Perception* (β= 0.620; t= 7.209; p= 0.000 [0.449, 0.791]; *Profile* (β= 0.326; t= 3.268; p= 0.001 [0.128, 0.524]; *Fluency* (β=-0.285; t= -2.587; p=0.011 [-0.504, -0.066]; *Form* (β= 0.277; t= 1.998; p= 0.049 [0.002, 0.553]; *Time-lag* (β= 0.125; t= 1.967; p= 0.0529 [-0.001, 0.251].

The 5 dimensions were integrated in a new dichotomized variable to be included in the moderation analysis to generate a theoretical model. To do this, a new variable

evaluation was created as the mean of these 5 dimensions and then dichotomized (Lower/Higher), taking into account the median of the new variable.

A linear regression analysis (see Table 5) was run to test the influence of W (*status of the evaluator)*, X (*evaluation*) and WX (*interaction*) on Y (*overall quality*).

Model	coeff β	SE	Z	*p*	LLCI	ULCI
Constant	0.7077	0.3098	2.284	0.0224	0.1004	1.3149
Status	1.6132	0.6197	2.6033	0.0092	0.3986	2.8277
Evaluation	2.594	0.6197	4.1861	0.0000	1.3795	3.8085
Interaction	2.4909	1.2394	2.0098	0.0444	0.0618	4.9200

Table 5. Linear regression analysis

The statistically significant results in β1, β2, and for the most part in β3, confirm the presence of a moderation effect.

β1 (Y: Evaluation) = 2.594 p=0.000: The *evaluation* of characteristics of the interpreter and his/her interpretation influences the perception of the *overall quality* of the interpreter's performance.

β2 (W: Status) = 1.613 p=0.009: The *status* of the evaluator (layperson versus semi-expert) influences the perception of the *overall quality* of the interpreter's performance.

β3 (WX: Interaction) = 2.491 p=0.044: The *status* (layperson versus semi-expert) intervenes in the effect of the *evaluation* on the perception of the *overall quality* of an interpreter's performance.

The results obtained (see Figure 2 and Table 6) show a conditional effect of the *evaluation* on the perception of *overall quality* in both status groups (laypersons versus semi-experts) β (Laypersons -0.5) = 1.349, p=0.013; β (Semi-experts -0.5) = 3.839, p=0.001.

Status	Effect	SE	Z	*P*	LLCI	ULCI
-0.5	1.3486	0.5435	2.4812	0.013	0.2833	2.4138
0.5	3.839	1.1138	3.4471	0.001	1.6564	6.0225

Table 6. Conditional effects of the focal predictor according to moderator values

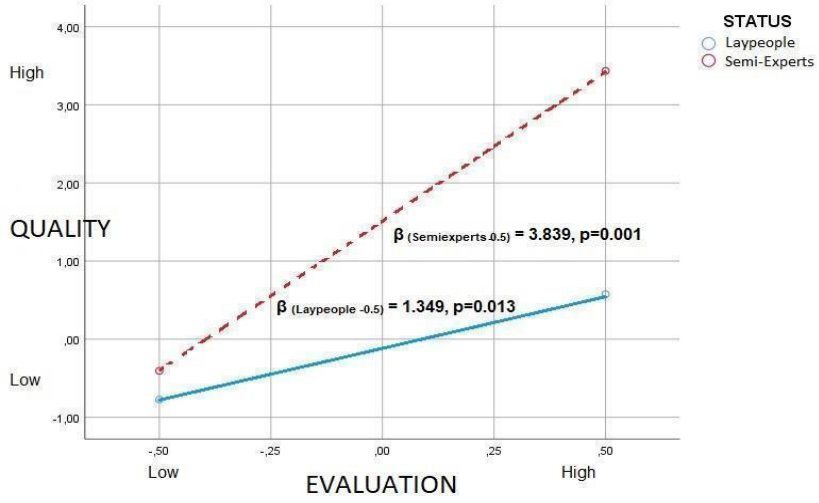


Figure 16: Effect of *overall quality* on status groups

4. Discussion and conclusions

4.1 Quality expectations

The criteria in which the semi-experts have higher expectations than the laypersons correspond to the innate and acquired skills of an interpreting professional, and also encompass aspects related to what may be called 'intrinsic mechanism' (Sánchez Santa-Bárbara/Pradas Macías 2019), referring to variables related to the interpreting process itself.

The one expectation criterion in which the laypersons have a significantly higher score than the semi-experts refer to the belief that interpretations have a monotonous intonation. Moreover, it appears that this expectation was confirmed when they listened to the SI and that they penalized this aspect in their evaluation, even though it formed part of their expectations.

With regard to the differences between expectations of men and women, women have higher expectations regarding the structure of the interpretation, expecting it to be well structured. They also expect to a greater degree than men that the interpreter will adapt to the original speaker.

4.2 Quality evaluation

4.2.1 Regression analysis

The regression analyses enabled us to identify which dimensions function better as predictors of the overall quality. In line with other previous research (see section 2),

aspects related to the evaluation of the delivery and professionality were considered less important than *perception*, *profile*, *fluency*, *form* and *time-lag*.

The analysis shows that the criteria *perception*, *form, fluency*, *presentation* and *time-lag* predict 68% of the overall perceived quality. This may be regarded as statistical proof that formal aspects have a salient impact on quality perception, as has been observed in previous studies with a two-pronged approach (*e.g*. Collados Aís 1998; Collados Aís *et al*. 2007). Indeed, the results in this study show that formal variables carry more weight than content variables, and even seem to dominate in the quality evaluation of a SI. What is more, fluency seems to play a determinant role. These results are highly interesting in the light of the discussion in different research fields about whether the criterion *fluency* should be associated with formal presentation or content (see section 2).

Researchers such as Pradas Macías (2003) and Rennert (2010, 2019) have specifically observed the impact of pauses (whether silent or with sound) on fluency and quality perception. The results of the present study point to an understanding of fluency as a dual-nature criterion, instead of assigning it to either the formal or content category.

4.2.2 Means analysis

Regarding the evaluation of *profile*, men consider the female interpreter better at communicating ideas effectively. This result, together with the results obtained for *overall quality* in the case of the female interpreter's performance and *difficult to understand / easy to understand* in the case of the male interpreter's performance, suggests a possible gender-related difference in the evaluation.

Semi-experts consider the interpreters to have high competence, while laypersons evaluate the competence of both interpreters less favourably than the semi-experts do.

A further issue to consider is that the semi-experts consider the interpreters' voice more pleasant than laypersons do, and the latter consistently evaluate the time-lag worse than semi-experts do. The original interpretation selected by the researchers for the study is consistent with the intonation denominated by Shlesinger (1994) *sui generis intonation*, with the characteristic time-lag or silence pattern of SI (Pradas Macías 2015). A possible explanation, therefore, is that laypersons associate the criterion *voice* with *intonation*, while semi-experts do not (see Fernández Iglesias 2007).

This explanation is reinforced by the statistical difference between semi-experts and laypersons in the evaluation of native-like pronunciation. Nearly 100% of semi-experts identify the interpreter's pronunciation as native. Furthermore, semi-experts clearly rate *fluency* more favourably than laypersons. Despite the fact that laypersons evaluate the interpretation better with regard to coherence and find it easier to follow, these aspects lead them to make a worse evaluation of the overall quality of the interpretation.

4.3 Quality perception depending on status

The moderation analysis confirmed that the IT held by people with more or less knowledge about interpreting have a significant impact on their evaluation of the quality of the interpretation they had listened to. According to this result, the higher or lower perception of the overall quality of the interpretation not only depends on the higher or lower evaluation of the interpreter's characteristics and performance, but is also moderated by the status of the evaluators, both in the case of the laypersons and the semi-experts. In other words, if the evaluation of these aspects is considered lower or higher (below or above the median value), it will produce a different overall quality affected not only by the previous evaluation of the specific aspects, but also by the status of the person as someone with lower or higher expertise in the field. This finding is consistent with results in other domains in which IT have been studied (see section 2) and provides support for applying IT to the area of interpreting quality.

4.4 Operationalization of IT

Correlating expectations with overall quality shows that the IT of semi-experts include aspects that are basic to interpretation training.

Regarding laypersons, bearing in mind that IT are acquired through experience in a particular setting (see 3.2.) and, that interpreters do indeed seem to have a *sui generis intonation* (Shlesinger 1994, see 5.2.2), intonation seems to be a clearly established criterion in laypersons' IT regarding interpreters.

4.5 Final remarks

As with any research, this study has some limitations. First, the characteristics of the sample (convenience and non-probabilistic, as well as the reduced size in both the laypersons and the semi-experts group) reduce the external validity of the study.

Our intention in future studies is to access 3 subsamples: experts (professional interpreters), semi-experts (interpreter trainees) and laypersons (potential users of interpreting services), and to ensure that all 3 subsamples are of a larger size. This would enable us to test our hypotheses about expectations, gender and IT more deeply and thus further improve our understanding of the perception of quality.

Nevertheless, our results also lead us to question the assumption that professional interpreters should be considered "experts" in the perception of quality in interpreting. Professional interpreters rarely use interpreting services, so the "experts" might well be non-interpreters who have considerable experience in listening to experienced interpreters. We have seen that interpreting quality expectations differ depending on the IT of evaluators, but it is indeed interesting that "laypersons" believe that interpreters have a monotonous intonation. There may thus be a need to clarify what the subjects in quality studies are experts in: performing interpreting services or using interpreting services? In this respect, perhaps a finer-grained profile of subjects than "experts" and "laypersons" would be desirable.

A further aim for future research is to explore the relationship between expectations, evaluation, gender of the interpreter and the evaluator, IT of quality in interpreting and the perception of quality.

Our objective in this respect is to generate a moderated mediation model which would include not only, as in this case, the participants' status or evaluation, in order to explain the quality of the interpreter and his or her work; but would also take into account variables related to the expectations and the gender of the interpreter and evaluator.

References

Baer B. J. / Massardier-Kenney, F. (2015) "Gender and Sexuality", in C. Angelelli / B. J. Baer (eds) *Researching Translation and Interpreting*, London/New York, Routledge, 83-96.

Bühler H. (1986) "Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters", *Multilingua* 5/4, 231-236.

Cámara Aguilera E. / Pradas Macías, E. M. (2019) "Implicaciones del género en el discurso interpretado", *Entreculturas* 10, 301-320.

Chiaro D. / Nocella G. (2004) "Interpreters' perception of linguistic and non-linguistic factors affecting quality: A survey through the World Wide Web", *Meta* 49/2, 278-293.

Chodkiewicz M. (2020) "Changes in undergraduate students' conceptual knowledge of translation during the first years of translation/interpreting- and foreign language-related education", *Perspectives* 10, 1-20.

Christen y Gracia L.M. (2020) "Impact of voice gender on the perception of simultaneous interpreting quality", in R. Barranco-Droege (ed.) *Solving the Riddle of Interpreting Quality. Dimensions and Challenges*, Granada, Comares, 25-58.

Collados Aís Á. (1998) *La evaluación de la calidad en interpretación simultánea: la importancia de la comunicación no verbal*, Granada, Comares.

Collados Aís Á. / Pradas Macías E.M. / Stévaux E. / García Becerra O. (eds) (2007) *La evaluación de la calidad en interpretación simultánea: Parámetros de incidencia*, Granada, Comares.

Collados Aís Á. / Iglesias Fernández E. / Pradas Macías E. M. / Stévaux E. (eds) (2011) *Qualitätsparameter beim Simultandolmetschen*, Tübingen, Narr.

Defrancq B. / Collard C. / Magnifico C. / Iglesias Fernández E. (2021) "Sex and gender in conference interpreting", in M. Albl-Mikasa / E. Tiselius (eds) *The Routledge Handbook of Conference Interpreting*, London/New York, Routledge, 414-427.

Du B. (J.) (2020) "Gender and interpreting; an overview and case study of a woman interpreter's media representation", in L. Von Flotow / H. Kamal (eds) *The Routledge Handbook of Translation, Feminism and Gender*, London/New York, Routledge, 159-170.

Dweck C.S. (2012) “Implicit theories”, in P. Van Lange / A. Kruglanski / T. Higgins (eds) *The Handbook of Theories of Social Psychology*, London, SAGE, 43-62.
Dweck C.S. / Elliott E.S. (1983) “Achievement motivation”, in P. H. Mussen / E. M. Hetherington (eds) *Handbook of Child Psychology Volume 4*, New York, Wiley, 643-691.
Dweck C.S. / Chiu C.Y. / Hong Y.Y. (1995) “Implicit theories: Elaboration and extension of the model”, *Psychological Inquiry*, 6/4, 322-333.
Dweck C.S. / Hong Y.Y. / Chiu C.Y. (1993) “Implicit theories individual differences in the likelihood and meaning of dispositional inference”, *Personality and Social Psychology Bulletin* 19/5, 644-656.
ECIS Research Group (2003) *Corpus ECIS 2003*. [Unpublished database], Granada, University of Granada.
Enríquez Raído V. (2018) “Teacher motivation and emotions vis-à-vis students’ positive perceptions of effective teaching and learning”, *Translation, Cognition & Behavior* 1/2, 361-390.
Fernández Iglesias E. (2007) “La incidencia del parámetro agradabilidad de la voz”, in Á. Collados Aís / E.M. Pradas Macías / E. Stévaux / O. García Becerra (eds) *La evaluación de la calidad en interpretación simultánea. Parámetros de incidencia*, Granada, Comares, 37-51.
García Becerra O. (2012) *La incidencia de las primeras impresiones en la evaluación de la calidad de la interpretación simultánea: un estudio empírico*, PhD Thesis, University of Granada.
García C. / Tapias Merino D. (2000) “La frecuencia fundamental de la voz y sus efectos en el reconocimiento de habla continua”, *Procesamiento del lenguaje natural* 26, 163-168.
Glück J. / Strasser I. / Bluck S. (2009) “Gender differences in implicit theories of wisdom”, *Human Development* 6/1, 27-44.
Hayes A.F. (2013) *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
Kelly G. (1955) *The Psychology of Personal Constructs*. New York, Norton.
Kurz I. (1993) “Conference interpretation: Expectations of different user groups”, *The Interpreters’ Newsletter* 5, 13-21.
López Varela A. (2016) “Gender issues during simultaneous interpreting in violence against women”, *Polissema* 16, 53-66.
Magnifico C. / Defrancq B. (2015) “Norms and gender in simultaneous interpreting: A study of connective markers”, *Translation & Interpreting* 12, 1-17.
Pöchhacker, F. (2013) “Researching quality: A two-pronged approach”, in O. García Becerra / E.M. Pradas Macías / R. Barranco-Droege (eds) *Quality in interpreting: Widening the scope*, Granada, Comares, 33-56.
Pradas Macías E.M. (2003) *Repercusión del parámetro pausas silenciosas en la fluidez,* unpublished PhD Thesis, University of Granada.
Pradas Macías E.M. (2015) “User awareness of sui generis pause patterns in simultaneous interpreting”, in C. Zwischenberger / M. Behr (eds) *Interpreting Quality: A Look Around and Ahead*, Berlin, Frank & Timme, 67-88.
Pradas Macías E.M. (2019) “Towards a comprehensive notion of fluency in simultaneous interpreting: An interdisciplinary approach”, in L. Degand / G.

Gilkin / L. Meurant / A.C. Simon (eds) *Fluency and Disfluency across Languages and Language Varieties*, Louvaine-La-Neuve, Presses Universitaires de Louvain, 139-158.
Rennert S. (2010) "The impact of fluency on the subjective assessment of interpreting quality", *The Interpreters' Newsletter* 15, 101-115.
Rennert S. (2019) *Redeflüssigkeit und Dolmetschqualität*. *Wirkung und Bewertung*, Tübingen, Narr Francke Attempto.
Roach E. (1978) "Principle in categorization", in E. Roach / B.B. Lloyd (eds) *Cognition and Categorization*, Hillsdale, NJ, Erlbaum, 28-48.
Russo M. (2016) "Orality and gender: A corpus-based study on lexical patterns in simultaneous interpreting", *MonTI* 3, 307-322.
Ryan S. / Mercer S. (2012) "Implicit theories: Language learning mindsets", in S. Mercer / S. Ryan / M. Williams (eds) *Psychology for Language Learning*, London, Palgrave MacMillan, 74-89.
Sánchez Santa-Bárbara E. / Pradas Macías E.M. (2019) "Las teorías implícitas en la evaluación de una interpretación simultánea: El sexo del intérprete como factor predictivo", *Entreculturas* 10, 341-351.
Shlesinger M. (1994) "Intonation in the production and perception of simultaneous interpretation", in S. Lambert / B. Moser-Mercer (eds) *Bridging the Gap. Empirical Research in Simultaneous Interpretation*, Amsterdam/Philadelphia, John Benjamins, 225-236.
Shlesinger M. *et al*. (1997) "Quality in simultaneous interpreting", in Y. Gambier / D. Gile / C. Taylor (eds) *Conference Interpreting: Current Trends in Research*, Amsterdam/Philadelphia, John Benjamins, 123-130.
Xiaoshu Y. /Yunzhang S. /Hongyu W. (2021) "Investigating effects of gender identity on interpreters' styles: A case study of multidimensional analysis", in H. Kaibou *et al*. (eds) *Proceedings of the 35th Pacific Asia Conference on Language, Information and Computation*, Shanghai, Association for Computational Linguistics, 463-475.
Zwischenberger C. (2010) "Quality criteria in simultaneous interpreting: An international vs. a national view", *The Interpreters' Newsletter* 15, 127-142.
Zwischenberger C. (2013) *Qualität und Rollenbilder beim simultanen Konferenzdolmetschen*, Berlin, Frank&Timme.

Appendix

Appendix 1. Expectations results by group and gender

	Status	Mean	t	df	*p*	sig
Interpreters are likely to be confident	Semi-experts	6.07	1.529	114	0.129	ns
	Laypersons	5.77				
Interpreters are likely to be competent	Semi-experts	**6.65**	4.341	94.366	0.000	***
	Laypersons	5.94				
Interpreters are likely to be organized	Semi-experts	**6.31**	2.256	113.255	0.026	*
	Laypersons	5.89				
Interpreters are likely to be reliable	Semi-experts	**6.35**	2.569	114	0.011	*
	Laypersons	5.87				
Interpreters are likely to be flexible	Semi-experts	**6.21**	3.333	113	0.001	**
	Laypersons	5.48				
Interpreters are likely to be calm	Semi-experts	**4.96**	2.144	113	0.034	*
	Laypersons	5.52				
Interpreters are likely to be good at communicating	Semi-experts	6.59	0.819	114	0.415	ns
	Laypersons	6.47				
Interpreters are likely to be neutral	Semi-experts	6.08	1.870	109.784	0.064	ns
	Laypersons	5.60				
Interpreters are likely to be professional	Semi-experts	6.63	1.288	114	0.200	ns
	Laypersons	6.42				
Interpreters are likely to be pleasant	Semi-experts	5.52	0.345	112	0.730	ns
	Laypersons	5.44				
Interpreters are likely to be resolute	Semi-experts	**6.57**	3.475	111.770	0.001	**
	Laypersons	5.94				
Interpreters are likely to be audacious	Semi-experts	**6.29**	3.466	107.641	0.001	**
	Laypersons	5.52				
Interpreters are likely to be expressive	Semi-experts	5.80	0.925	114	0.357	ns
	Laypersons	6.02				
Interpreters are likely to have a pleasant voice	Semi-experts	5.28	1.323	113	0.189	ns
	Laypersons	5.60				
Interpreters are likely to have specialized knowledge about the topic	Semi-experts	5.83	1.874	113	0.064	ns
	Laypersons	5.55				
Interpreters are likely to have a monotonous intonation	Semi-experts	2.61	4.109	112.194	0.000	***
	Laypersons	**3.89**				

	Status	Mean	t	df	*p*	sig
Interpreters are likely to have specialized vocabulary	Semi-experts	**5.61**	2.486	114	0.014	*
	Laypersons	5.00				
Interpreters are likely to be calm under pressure	Semi-experts	6.20	0.333	113	0.740	ns
	Laypersons	5.73				
Interpreters are likely to have native-like pronunciation	Semi-experts	4.60	0.333	113	0.740	ns
	Laypersons	4.52				
Interpreters are likely to vocalize clearly	Semi-experts	6.37	1.171	114	0.244	ns
	Laypersons	6.16				
Interpreters are likely to speak fast	Semi-experts	3.81	0.361	112	0.719	ns
	Laypersons	3.92				
Interpreters are likely to speak similarly to the original speaker	Semi-experts	5.19	0.134	113	0.894	ns
	Laypersons	5.23				
Interpreters are likely to use formal register	Semi-experts	5.21	1.540	113	0.126	ns
	Laypersons	4.82				
Interpreters are likely to speak correctly	Semi-experts	**6.24**	2.380	114	0.019	*
	Laypersons	5.81				
Interpreters are likely to go unnoticed	Semi-experts	**5.17**	3.235	112	0.002	**
	Laypersons	4.16				
Interpreters are likely to adapt to the user	Semi-experts	**6.32**	2.723	106.158	0.008	**
	Laypersons	5.74				
Interpreters are likely to show confidence	Semi-experts	**6.37**	2.679	114	0.008	**
	Laypersons	5.90				
The interpretation is fluent	Semi-experts	**6.30**	2.232	114	0.028	*
	Laypersons	5.92				
The speed of the interpretation is fast	Semi-experts	4.17	1.097	114.000	0.275	ns
	Laypersons	4.48				
The rhythm of the interpretation is similar to that of the OS	Semi-experts	5.83	1.589	112.167	0.115	ns
	Laypersons	5.48				
The speed of the interpretation is similar to that of the OS	Semi-experts	5.61	0.563	114.000	0.575	ns
	Laypersons	5.48				
The interpretation has speed variations	Semi-experts	4.74	0.272	103.420	0.787	ns
	Laypersons	4.82				
The interpretation is faithful to OS	Semi-experts	**6.22**	2.811	105.967	0.006	**
	Laypersons	5.56				
The interpretation is complete	Semi-experts	5.47	0.542	113.000	0.589	ns
	Laypersons	5.60				

	Status	Mean	t	df	*p*	sig
The interpretation has few pauses	Semi-experts	4.92	1.146	112.000	0.254	ns
	Laypersons	4.58				
The interpretation is well structured	Semi-experts	6.22	0.858	114.000	0.393	ns
	Laypersons	6.06				
The interpretation has inappropriate pauses	Semi-experts	2.75	0.821	112.000	0.413	ns
	Laypersons	3.02				
The interpretation has incomplete sentences	Semi-experts	2.13	1.849	111.286	0.067	ns
	Laypersons	2.65				
The interpretation has long pauses	Semi-experts	2.56	1.226	112.000	0.223	ns
	Laypersons	2.92				
The interpretation is clear	Semi-experts	**6.54**	2.359	91.662	0.020	*
	Laypersons	6.08				
The interpretation is easy to understand	Semi-experts	6.50	3.474	89.805	0.001	**
	Laypersons	**5.81**				

* p<0.05; ** p<0.01 *** p<0.001

	Gender	Mean	t	df	*p*	sig
Interpreters are likely to be confident	Male	5.93	0.156	113.000	0.876	ns
	Female	5.90				
Interpreters are likely to be competent	Male	6.24	0.123	113.000	0.902	ns
	Female	6.27				
Interpreters are likely to be organized	Male	6.03	0.258	113.000	0.797	ns
	Female	6.09				
Interpreters are likely to be reliable	Male	6.10	0.099	113.000	0.921	ns
	Female	6.08				
Interpreters are likely to be flexible	Male	5.76	0.248	112.000	0.804	ns
	Female	5.82				
Interpreters are likely to be calm	Male	5.34	0.441	112.000	0.660	ns
	Female	5.21				
Interpreters are likely to be good at communicating	Male	6.38	1.082	113.000	0.281	ns
	Female	6.57				
Interpreters are likely to be neutral	Male	5.79	0.061	112.000	0.952	ns
	Female	5.81				
Interpreters are likely to be professional	Male	6.55	0.272	113.000	0.786	ns
	Female	6.50				

	Gender	Mean	t	df	*p*	sig
Interpreters are likely to be pleasant	Male	5.29	0.941	58.889	0.351	ns
	Female	5.52				
Interpreters are likely to be resolute	Male	5.93	1.749	112.000	0.083	ns
	Female	6.32				
Interpreters are likely to be audacious	Male	5.50	1.711	110.000	0.090	ns
	Female	5.98				
Interpreters are likely to be expressive	Male	5.69	1.047	113.000	0.297	ns
	Female	5.98				
Interpreters are likely to have a pleasant voice	Male	5.21	1.323	113.000	0.189	ns
	Female	5.52				
Interpreters are likely to have specialized knowledge about the topic	Male	5.28	1.874	113.000	0.064	ns
	Female	5.80				
Interpreters are likely to have a monotonous intonation	Male	3.55	0.908	113.000	0.366	ns
	Female	3.20				
Interpreters are likely to have specialized vocabulary	Male	5.03	1.088	113.000	0.279	ns
	Female	5.35				
Interpreters are likely to be calm under pressure	Male	**5.64**	1.759	110.000	0.081	ns
	Female	6.02				
Interpreters are likely to have native-like pronunciation	Male	4.97	1.808	112.000	0.073	ns
	Female	4.42				
The interpreters vocalise clearly	Male	6.28	0.153	113.000	0.879	ns
	Female	6.24				
Interpreters are likely to speak fast	Male	4.28	1.600	111.000	0.113	ns
	Female	3.71				
Interpreters are likely to speak similarly to the original speaker	Male	4.90	1.257	112.000	0.211	ns
	Female	5.29				
Interpreters are likely to use formal register	Male	5.03	0.242	112.000	0.809	ns
	Female	4.96				
Interpreters are likely to speak correctly	Male	5.86	0.858	113.000	0.393	ns
	Female	6.05				
Interpreters are likely to go unnoticed	Male	4.21	1.461	111.000	0.147	ns
	Female	4.75				
Interpreters are likely to adapt to the user	Male	5.45	2.434	36.863	0.020	*
	Female	6.19				
Interpreters are likely to show confidence	Male	**5.76**	1.883	35.896	0.068	ns
	Female	6.23				

	Gender	Mean	t	df	*p*	sig
The interpretation is fluent	Male	5.83	1.766	113.000	0.080	ns
	Female	6.17				
The speed of the interpretation is fast	Male	4.72	1.625	113.000	0.107	ns
	Female	4.19				
The rhythm of the interpretation is similar to that of the OS	Male	5.38	1.320	113.000	0.189	ns
	Female	5.72				
The speed of the interpretation is similar to that of the OS	Male	5.62	0.463	113.000	0.644	ns
	Female	5.50				
The interpretation has speed variations	Male	4.76	0.059	113.000	0.953	ns
	Female	4.78				
The interpretation is faithful to OS	Male	5.72	0.638	113.000	0.525	ns
	Female	5.91				
The interpretation is complete	Male	5.66	0.653	112.000	0.515	ns
	Female	5.48				
The interpretation has few pauses	Male	4.86	0.533	111.000	0.595	ns
	Female	4.68				
The interpretation is well structured	Male	5.69	2.278	35.401	0.029	*
	Female	6.28				
The interpretation has inappropriate pauses	Male	**3.24**	1.201	111.000	0.232	ns
	Female	2.80				
The interpretation has incomplete sentences	Male	2.86	1.811	113.000	0.073	ns
	Female	2.27				
The interpretation has long pauses	Male	3.14	1.501	111.000	0.136	ns
	Female	2.63				
The interpretation is clear	Male	5.97	1.828	113.000	0.070	ns
	Female	6.40				
The interpretation is easy to understand	Male	5.86	1.387	113.000	0.168	ns
	Female	6.21				

* $p<0.05$; ** $p<0.01$ *** $p<0.001$