

Communication mediated by conference interpreters: age and sex stereotypes

Estereotipos relativos a la edad y el sexo aplicados a la comunicación mediada por intérpretes de conferencias

7

ARTICLE



Lucila Christen y Gracia

Universitat de Vic- Universitat Central de Catalunya

Lucila Christen y Gracia has been a professional interpreter and translator since 1973. She holds a Ph.D. in Translation, Gender, and Cultural Studies (*cum laude*) from the University of Vic-Central University of Catalonia, Spain, where she obtained a Master's degree in specialized legal-financial translation (Extraordinary Prize). She has been a member since 1982 and former president of the Mexican Association of Conference Interpreters (CMIC), an Honorary Member of the Mexican Association of Graduates in Translation and Interpreting (CMLTI), and a member of the American Translators Association (ATA). Since 2002, she has been President and CEO of Grupo Gaute, the first ISO-certified translation and interpreting agency in Latin America.

lucilamaria.christen@uvic.cat

ORCID: <https://orcid.org/0000-0002-8531-0870>

RECEIVED: 2023-02-25 / ACCEPTED: 2023-04-24

Abstract

Previous studies have concluded that the combination of interpreters' sex and age may affect the perception of their simultaneous interpretation. In this study, 156 subjects completed a questionnaire assessing various non-verbal factors of the simultaneous interpretations produced by four male and four female interpreters of different age groups in a recording studio. The recordings were controlled to

achieve a homogeneous rendition of verbal factors to favor the rating of non-verbal factors. The results show that interpreters' sex- and age-related characteristics, as discerned by the listener, may bias the perceived interpreter's performance.

KEYWORDS

Sex and age, stereotypes, ageism, simultaneous interpretation.

Resumen

Estudios previos concluyen que la combinación del sexo y la edad de los y las intérpretes puede condicionar la percepción de la interpretación simultánea. En este estudio, 156 participantes contestaron un cuestionario para evaluar diversos factores no verbales de las interpretaciones realizadas por ocho intérpretes, cuatro hombres y cuatro mujeres, grabadas en un estudio. Las grabaciones se controlaron para lograr una producción homogénea de los fac-

tores verbales que no interfiriera en la evaluación de los factores no verbales. Los resultados muestran que las características relacionadas con el sexo y la edad de los y las intérpretes pueden sesgar la percepción de la interpretación simultánea por parte de la audiencia.

PALABRAS CLAVE

Sexo y edad, estereotipos, edadismo, interpretación simultánea.

1. INTRODUCTION AND BACKGROUND

This research builds upon a previous Baseline Research (Christen, 2020), which investigated the impact of interpreter sex on the perception of simultaneous interpreting (SI). However, the previous research faced challenges with a small sample size (Gile, 2018), so the authors made the methodological improvements outlined in the methodology section of this publication. This study aims to replicate the Baseline Research and either support or refute its findings. The present study will consider the age and sex of the interpreters, as suggested in the conclusions of the previous research. For this research, sex will be used dichotomously (Bhargava et al., 2021) without including social genders. The study included interpreters with no borderline voices, i.e., no male voices in the female frequency range (135 Hz to 635 Hz) and no female voices in the male frequency range (75 Hz to 480 Hz). This study is not intended to discriminate based on age, sex, ethnicity, or any other factor, nor to promote stereotypes that might lead to such exclusions. When referring to interpreters or participants, men and

women are included unless the sex of the specific group under study is specified for clarity. Where the context of the text so requires, a distinction between sexes may be made.

1. 1 INTERPRETATION STUDIES

Simultaneous interpreting is essential for multilingual communication at congresses and events. With the help of technological tools, the interpreter's oral production is delivered synchronously with the original speech, delayed by the time it takes the professional to process it, creating a sense of simultaneity for the listener.

There is a lack of common ground in SI studies when investigating the sex (Defrancq et al., 2021) and the age of interpreters. Most recent studies have focused on investigating the impact of sex differences in interpreters on the one hand and the age or experience of interpreters on the other.

Regarding the sex of the interpreter, several topics have been addressed, such as the role of sex and linguistic politeness in police interviews (Nakane, 2008) or sex differences in facial

threat and the use of mitigating factors (hedg-
es) in the treatment of politeness in SI in the
European Parliament (Magnifico, 2017). Some
studies have focused on the influence of the
sex of the interpreter on the interpreting lag
(Ear-Voice-Span), showing that the length of
the lag varies significantly between interpreters
(Collard & Defrancq, 2017, 2019). A study
that also deals with the sex of the interpreters
concludes that connectors vary according to in-
terpreting norms and the sex of the interpreter
(Magnifico & Defrancq, 2020). Finally, according
to Bartłomiejczyk (2020), female interpreters
mitigate serious and deliberate rudeness to a
greater extent than male interpreters.

There is less research on the age or experience
of the interpreter. Liu (2008) concludes that the
main differences between novice and experi-
enced interpreters lie in the strategies the lat-
ter use in comprehension, translation, and pro-
duction and in their ability to switch between
these processes.

Along the same line, Tiseliu (2013) analyses
the process and product of interpreting in in-
terpreters with different levels of experience
and concludes that interpreters with long pro-
fessional experience show significant differ-
ences in their interpreting skills compared to
those with little or no interpreting experience.
According to the results of her study, the for-
mer find fewer differences in speech process-
ing and have more problem-solving strategies.
Subsequently, Liu et al. (2020) interviewed in-
terpreters aged 70 and over who are members
of AIIC (Association Internationale des Inter-
prètes de Conférence).

Most respondents acknowledge that they have
faced some challenges in the later stages of
their careers, which they attribute to changes
in the working environment. Chmiel (2021) ex-
amines the latency of word interpretation and
points out that it improves during interpreter
training but does not increase with profession-
al experience. Moser-Mercer (2022) suggests
that SI experience should be analyzed not only
in terms of behavioral and brain changes but
also in terms of the needs of the profession.

Regarding the inclusion of age and sex, Collados
Aís (2007), from a different approach, examines
users' perceptions of the quality of interpreting
according to the sex and age of the user. Few
studies include the age and sex of the interpret-
ers in their approach. Angelelli (2004) includes
age, sex, level of education, and income when
studying the role of the interpreter in commu-
nity, legal, and medical settings to establish the
relationship between these variables and the
self-perceived visibility of interpreters.

According to the results of this study, there
were no significant differences between the
groups investigated. Regarding age, the older
participants perceived themselves as less visi-
ble than the younger ones. García Becerra, who
studied the effect of first impressions on the
perception of the quality of SI, points out that
“[...] subjects recognized that their expectations
could vary according to certain conditions or
characteristics of the interpreter, such as age,
sex or vocal aspects” (2012, p. 566).

In a later study analyzing the effect of order and
first impressions on the perception of SI, García

Becerra (2015) points out that in-depth studies related to first impressions as a function of the interpreter's sex and age are needed. In the evaluations carried out by the interviewees, negative traits had a negative influence on the interpreters' perception of their performance. Conversely, positive traits improved the interpreters' evaluations more than those of their female colleagues. Drawing on theories from the sociology of professions, Gentile (2016) investigates the self-perceived professional status of conference and public service interpreters. Her questionnaires (one for each type of interpreting) include, among others, demographic elements, including sex and age variables. It concludes that the increasing feminization of the profession, rapid technological change, and a complex labor market, along with other factors, have influenced interpreters' perceptions of the profession.

1.2 EXTERNAL DISCIPLINES

To make sense of the world, humans need to categorize their perceptions and group them into objects and events of a similar nature, identify shared characteristics, and reduce the time needed to process redundant information (Cuddy & Fiske, 2004). Just as we group objects and events, we group people based on perceived similarities. In our eagerness to assign objects, events, and people to meaningful classes about which we have established beliefs and expectations lies the root of stereotypes (Cuddy & Fiske, 2004). These variables are autonomous in their motivations and arise from the activation of patterns learned in each person's life cycle. Humans can decode the information contained

in verbal messages on several levels (Nass & Brave, 2005). Fundamental frequency (FO) and formant frequencies, which are the most important vocal features for identifying the sex and age of an individual, can be extracted from speech signals and encoded almost immediately in brain areas specialized in the perception of male or female voices. (Abdulsatar et al., 2019; Zhang et al., 2020). As a result, listeners easily develop stereotypical responses to the sex and age of speakers based on the listeners cultural norms (Pisanski & Feinberg, 2015). Age and sex stereotypes automatically interact to influence the processing of speech perception (Strand, 2000).

It is widely known that age stereotypes precede age discrimination. Voss et al. (2018) point out that a reference to age does not equate to age discrimination (ageism). Age categorization becomes ageist only when described as a character based on stereotypes. To evoke an ageist stereotype, several elements of judgment must interact and, almost always unconsciously, the factor with the most substantial impact is the voice of the older adult (Hummert et al., 1999). The content and activation of sex stereotypes reflect a bias that favors men over women (Casper & Rothermund, 2012; Kornadt et al., 2013; Krekula et al., 2018).

Given the above, the following questions arise: Can the sex, and age of the interpreter bias the perception of SI? Could this bias be due to cultural norms and stereotypes of listeners derived from the physiological characteristics they unconsciously perceive in the speaker?

2. METHODOLOGY

In this quantitative quasi-experimental study, a closed-ended questionnaire was applied to corroborate the findings of the Baseline Research without modifying its methodology in terms of the age and sex of the interpreters, the questionnaire, the interpreted performances, the recording environment, and the recruitment of participants. Improvements included an increased number of interpreters and respondents, a differentiated control of the recording groups, the evaluation of the linguistic criteria of the recorded performances, and a new statistical design to ensure the randomization of the recordings according to the sex and age of the interpreters.

2.1 INTERPRETERS

Given that users show a certain tendency to evaluate an SI with a non-native accent negatively (Cheung, 2020), native Spanish-speaking interpreters were chosen based on their place of birth and their academic background in Mexico City. The study included two groups. Each group consisted of two novice and two experienced interpreters, one male and one female per age group (see Table 1).

The first group (G1) participated in the Baseline Research, while the second group (G2) was recruited for this study. The four experienced interpreters of both sexes, aged 60 and over, are graduates with at least 25 years of continuous practice in IS. The four novice interpreters, male and female, aged 26-34, are graduates with at most ten years of continuous practice in SI. All eight interpreted their renderings from English into Spanish.

Table 1

Age of interpreters at the time of the study and years of experience

Group	Code	Age on date of recording	Years of uninterrupted practice
1	G1EF	68	
1	G1EM	69	48
1	G1NF	26	1
1	G1NM	27	2
2	G2EF	69	42
2	G2EM	60	33
2	G2NF	27	6
2	G2NM	34	9

Note. G = group; M = female; H = male; E = experienced; N = novice.

2.2 PRESENTATIONS

Four videos in English were chosen from among the speeches that the global TED (Technology, Entertainment and Design) community disseminates through the Internet (see Table 2). These talks, approximately 15 minutes long, were selected to reduce the number of interferences that could negatively affect the primary sources of cognitive loads in the interpretation (Riccardi, 2022). Special attention was paid to the speaker's accent, intonation, and pronunciation (Pöchhacker, 2016), the excessive speed of the original speech (García et al., 2020), the complexity or density of the source speech, and the use of technical jargon or syntactically complex sentences (Gile, 2002).

2.3 SEQUENCE DESIGN

To provide greater methodological rigor, to make the results robust, and to ensure complete randomness in the order of the recording segments by avoiding repetition of the ages and sexes of the interpreters and the sex of the speakers in the sequences, two researchers from the Centro de Investigación en Matemáticas (CIMAT), in León, Guanajuato, Mexico, designed four thirty-minute sequences. Each sequence included the interpretation of two talks, one per group (G1 and G2), with eight segments of approximately three and a half minutes, one segment per interpreter. Subsequently, the order of the groups was alternated to obtain a total of 64 segment positions in eight sequences (see Table 3).

Table 2

Audiovisual material used in the experiment

Pre-sentation Name	Link	Name of speaker	Short name
"The gentle power of highly sensitive people"	https://www.youtube.com/watch?v=pi-4JOIMSWjo	Elena Herdieck-erhoff	Sensitive
"After watching this, your brain will not be the same again."	https://www.youtube.com/watch?v=L-NHBMF-CzznE	Lara Boyd	Brain
"What you didn't know about coffee."	https://www.youtube.com/watch?v=JaQNyOE-f4YY	Asher Yaron	Coffee
"How to become a millionaire in three years."	https://www.youtube.com/watch?v=-jvBaRf9L-HDs	Dany Ally	Millionaire

2.4 RECORDING PROCESS

A specialist with a master's degree in engineering and a recording engineer designed the recording process and implemented it in both groups. Using the 4K Downloader software (Open Media, LLC, Nizhny Novgorod, Russia), the four presentations were downloaded from the YouTube platform in the original 1280 x 720 format, with H.264 compression, in MP4. The audio and video source files were imported into a ProTools HD Ultimate™ session. Interpreters viewed the video on a screen and listened to the audio through their headphones. The performances were recorded in 48 kHz, 16-bit digital audio in WAV format through a DAKING Mic Pre-500 preamplifier with no analog compression and using SSL Alphalink audio interfaces.

The recordings had a target of approximately 28 LUFS and maximum peaks of -6 dB, which depended on the individual interpreter. All audios had a low-frequency cutoff at 80 Hz. The multiband normalization, compression, and limiting processes were performed with very subtle parameters to preserve the dynamics of the voices.

Adobe *Premiere* Pro™ was used to generate the final files. The MP4 source files and the WAV file with the voices of the interpreters were imported into a new *Premiere* session. In each video, the mastered audio was integrated with the original audio. In this integration, the original English audio was kept in the left channel, and the recorded Spanish audio was assigned to the right channel. To ensure synchrony, a time-code was used on both channels.

Table 3*Outline of the eight recording sequences designed for the study*

No.	Pres-entation	Speak-er	Segment				Segment				Speak-er	Pres-entation
			1	2	3	4	1	2	3	4		
			Group 1				Group 2					
1	Sensitive	FS	G1EF	G1NM	G1NF	G1EM	G2NF	G2EM	G2NM	G2EF	MS	Coffee
2	Millionaire	MS	G1NM	G1NF	G1EM	G1EF	G2NM	G2EF	G2NF	G2EM	FS	Brain
3	Coffee	MS	G1EM	G1NF	G1NM	G1NF	G2EM	G2NF	G2EF	G2NM	FS	Sensitive
4	Brain	FS	G1NF	G1EM	G1EF	G1NM	G2EF	G2NM	G2EM	G2NF	MS	Millionaire
			Group 2				Group 1					
5	Coffee	MS	G2NF	G2EM	G2NM	G2EF	G1EF	G1NM	G1NF	G1EM	FS	Sensitive
6	Brain	FS	G2NM	G2EF	G2NF	G2EM	G1NM	G1NF	G1EM	G1EF	MS	Millionaire
7	Sensitive	FS	G2EM	G2NF	G2EF	G2NM	G2EM	G1EF	G1NM	G1NF	MS	Coffee
8	Millionaire	MS	G2EF	G2NM	G2EM	G2NF	G1NF	G2EM	G1EF	G1NM	FS	Brain

Note. FS = female speaker; MS = male speaker; M = male; F = female; E = experienced; N = novice.

Source: Andrés Christen PH.D., Sebastián Quintanilla, M. Sc.

The G1 participants received the transcripts and glossaries of two talks (Millionaire and Coffee) one day in advance. Before starting the recording session, the interpreters reviewed the other two talks (Brain and Sensitive) for a few minutes. All group members recorded their four performances in their entirety in one day, at staggered times. The 16 segments used for this group's sequences were extracted from the recordings according to the position assigned to each interpreter in the designed sequence. Participants in G2 received the four transcripts of the talks and their glossaries two days in advance. Each interpreter was given a specific recording day. This group recorded only the segments that had been assigned to each interpreter during the design of the sequences.

This differential treatment of the G2 group in terms of preparation, time, and recording method was designed to enable G2 participants to produce performances with better control of the verbal features in their oral production than G1 participants. The aim was to assess whether respondents could discriminate such differences despite the randomized design of the sequences. If so, this design would provide a higher level of validity in assessing of the parameters under analysis.

To ensure that the interpretations of all participating interpreters, regardless of sex or experience, were perceived to be of a comparable linguistic and terminological level, the eight interpreters were provided with transcripts of

the four speeches during the recordings. They were reassured that the study aimed to obtain interpretations with the best possible linguistic characteristics, reflecting a complete and faithful translation of the original speech. They were also offered the opportunity to re-record their interpretations if they were not satisfied with the first recording.

2.5 QUESTIONNAIRE

The 5-point questionnaire used in the study was based on previous instruments on prosodic features and work attitudes tested for their relationship with perceived SI quality (Collados Aís *et al.*, 2007). It included five prosodic criteria: voice (Iglesias Fernández, 2007), intonation (Collados Aís, 2019), fluency (Rennert, 2020), diction (Blasco Mayor & García Becerra, 2007), and rhythm, the latter given its demonstrated importance in SI anticipation (Palová & Kiktová, 2019). It also included two work attitudes (credibility and professionalism) and one linguistic parameter (cohesion) (Collados Aís *et al.*, 2007). Although Collados Aís *et al.* (2007) studied cohesion, in the focus group conducted to verify the validity of this study, it was determined that the term coherence should be used instead due to cultural issues in the environment where the questionnaire would be applied. For this reason, as shown in Figure 1, coherence was the criterion to be assessed in the questionnaire. Finally, overall perceived quality was included to measure the user's satisfaction with the interpretation.

The same questionnaire was used for the Baseline Research and this study. Its validity and reliability were tested in a focus group mentioned above in which four men and four women from four age groups, one man and one woman per age group, were interviewed after viewing one of the recorded talks given by the two groups of interpreters. The validity of the questionnaire

was confirmed by the ability of the focus group participants to distinguish the perceived age and sex characteristics of the interpreters and to rate the interpretations. Once the questionnaire was designed (see Figure 1), the template was repeated eight times on individual sheets, one for each segment, to complete the measurement tool.

2.6 PROCEEDINGS

The minimum number of subjects (100) recommended to obtain meaningful results in quantitative studies (Dörnyei, 2007) was taken as a basis. One hundred and fifty-six native Spanish-speaking subjects, 79 male and 77 female, divided into five age groups, were recruited using a snowball sampling method and answered the questionnaire to rate the sequences recorded by the eight interpreters (see Table 4). The participants were aged 20 years or older, lived in the metropolitan area of Mexico City, and had at least one previous contact with professional SI services given their lack of knowledge of English, the language generally used in seminars and conferences in Mexico City. In total, 27 had a high school diploma, 45 were university students, 68 had a bachelor's degree, and 16 had completed postgraduate studies.

Table 4

Experimental sample by age range and sex

Sex	Age group				
	20 to 30	31 to 40	41 to 50	51 to 64	65 to 75
Female	10	23	18	18	8
Male	18	13	17	19	12
Total	28(17,9 %)	36(23,0%)	35(22,4%)	37(23,7%)	20(12,8%)

Note. n=156.

Figure 1*Outline of the questionnaire designed for the study*

Question	Rating scale				
1. Please rate the quality of the interpretation	<i>Very bad</i>				<i>Very good</i>
	1	2	3	4	5
2. Please rate the coherence in the transmission of the interpretation	<i>Very bad</i>				<i>Very good</i>
	1	2	3	4	5
3. Please rate the fluency of the interpretation	<i>Not fluency</i>				<i>Full fluency</i>
	1	2	3	4	5
4. Please rate the diction of the interpreter	<i>Not clear</i>				<i>Absolutely clear</i>
	1	2	3	4	5
5. Please rate the intonation of the interpreter	<i>Very monotonous</i>				<i>Not at all monotonous</i>
	1	2	3	4	5
6. Please rate the rhythm maintained by the interpreter	<i>Not regular at all</i>				<i>Very regular</i>
	1	2	3	4	5
7. Please rate the voice of the interpreter	<i>Not pleasant at all</i>				<i>Very pleasant</i>
	1	2	3	4	5
8. Please rate the credibility of the interpreter	<i>Not credible</i>				<i>Very credible</i>
	1	2	3	4	5
9. Please rate the professionalism of the interpreter	<i>Not professional</i>				<i>Very professional</i>
	1	2	3	4	5

The study was conducted in 41 sessions from 7 November to 12 December 2020. Each participant sat at an individual table and received disposable headphones, the eight-page questionnaire, and a pencil. The room was arranged to emulate a conference setting. Audio recordings of the TED talks in English were played through loudspeakers while participants simultaneously listened to the voices of the Spanish interpreters in their headphones.

Before the exercise, the facilitators explained to the participants the operationalized definitions of the parameters of the questionnaire and the written and oral instructions for completing the questionnaire. The facilitators played the video segments of the selected sequence one by one on a screen at the front of the room, stopping

the projection at the end of each segment. A response time of 90 seconds for each segment was given to ensure spontaneous reactions from participants to the voices of the interpreters (García Becerra, 2012; Larrea Estefanía, 2014). In each session, facilitators presented one of the eight sequences, reaching a total of 20-23 respondents per sequence. Participants did not have access to the interpreters' photographs or CVs, so their assessments were based solely on their impressions of the interpreters' voices (Horváth, 2017). The president and vice-president of the Colegio Mexicano de Intérpretes de Conferencias, A.C. 2020-2021, attended some sessions to verify full compliance with the designed method.

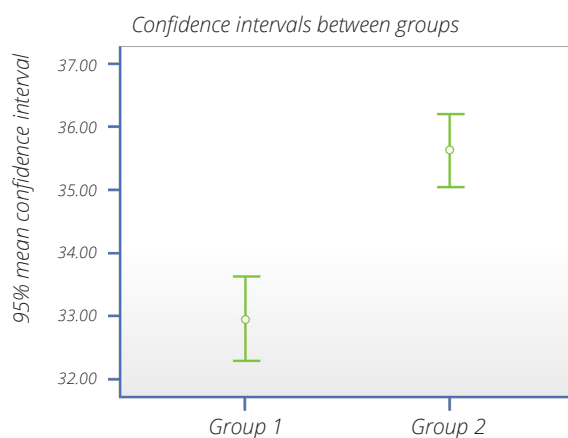
Table 5*Between-group ANOVA test*

Independent sample test										
		Levene's test for equality of variances		t-test for equality of means						
		F	Equal	t	df	Sig. (bi-lateral)	Mean difference	95 % confidence interval of the difference		
								Standard error difference	Lower	Upper
Average	Equal variances assumed	18.202	.000	5.938	1246	.000	-.29380	.04948	.39087	-.19673
	Equal variances not assumed			5.938	1219.291	.000	-.29380	.04948	.39088	-.19673

Note. F = ratio of the two variances; df = degrees of freedom

3. RESULTS

The first objective was to determine whether respondents perceived a significant difference in the rating of interpretations between G1 and G2. A one-way analysis of variance (ANOVA) test was performed to determine the mean difference between the groups (-0.2938). The result was significant, greater than 1%, in favor of G2, with a p-value of 0.000 (see Figure 2 and Table 5).

Figure 2

This result shows that respondents gave higher ratings to the G2 renditions than to the G1 renditions, even though the precise segmentation of the randomized sequential design could have confounded this overall perception. The results show that the longer preparation time given to the G2 interpreters to study the talks and the more favorable conditions for their recordings led to better control of the substantive features of their performances. In summary, as expected, the G2 interpreters working under more favorable conditions outperformed the G1 interpreters, confirming the robustness of the sequential design.

Tukey tests were used to compare the group means pairwise for each of the nine parameters. The null hypothesis was that the means were equal between groups. The results were significant at the 5% confidence level for the sample size ($n = 156$).

A large number of tests were conducted in this study, which could lead to Type I error inflation.

When and how to correct for multiple tests remains a hotly debated topic (Streiner, 2015). In this study, a non-confirmatory, albeit exploratory, post-hoc analysis was performed whose significance level rarely needs to be corrected, as observed in multiple regression analysis with categorical variables mathematically equivalent to ANOVA (Streiner, 2015).

The following pairs were analyzed:

- (1) Experienced female vs. experienced male (EF vs. EM).
- (2) Experienced female vs. novice female (EF vs. NM).
- (3) Experienced female vs. novice male (EF vs. NM)
- (4) Experienced male vs. novice female (EM vs. NF)
- (5) Experienced male vs. novice male (EM vs. NM)
- (6) Novice male vs. novice female (NM vs. NF)

When assessing perceived quality, the difference between ME means with the other three groups (EM, EF, and NF) were -0.349, -0.369, and -0.577, respectively, significant in favor of all three groups, with a confidence level of at least 5% in all cases. The difference between the means of EM and NF was -0.019, with a standard error of 0.077, the average between the two being approximately the same, so the null hypothesis cannot be rejected. The difference between the means of EM and NM was -0.228, with a confidence level of at least 5%. The comparison between the NM and NF groups gave an average of -.208 with a confidence level of 5%.

When examining coherence, significant differences were found between the means of EF and EM, with a difference of -0.218 and a con-

fidence level of at least 5%. Additionally, there were differences found between EF and NF, as well as between EF and NM, with differences of -0.333 and -0.462, respectively, and a confidence level of at least 5% in both comparisons. However, when comparing EM and NF, no significant difference was found, with a mean difference of -0.115 and a confidence interval that included zero. On the other hand, the mean difference between EM and NM was -0.244 in favor of the NM group, with a confidence level of at least 5%. Finally, when comparing NM and NF, no significant difference was found.

For fluency, the difference between the EF and EM means was -0.250, in favor of the EM group, significant at 5%. The difference between the EF and NF groups was -0.196, significant only at 10% and in favor of the NF group. The mean difference between EF and NM was -0.506, in favor of the NM group, with a confidence level of at least 5%. The comparison between the EM and NF groups was not significant. The mean difference between NM and EM was 0.256 in favor of the former, significant at 5%; between the NM and NF groups, the difference was 0.311 points in favor of the former, with a confidence level of at least 5%.

When assessing the diction of the interpreters, the difference between the EF and EM means was -0.151, in favor of the EM group; this value was not significant. The mean differences between EF and NF and between EF and NM were -0.330 and -0.484, respectively. These differences were significant in both comparisons, at least 5 %. The mean difference between EM and NF showed no significant difference between these groups. In the case of EM and NM, the main difference was -0.333 in favor of the NM group, at a confidence level of at least 5%. The comparison between NM and NF showed no significant difference.

The mean differences in intonation between EF and the other three groups (EM, NF, and NM) were significant, reaching -0.375, -0.417, and -0.394, respectively, in favor of the latter three, with a confidence level of at least 5% in all pairwise comparisons. The comparison between EM and NF showed no significant differences, suggesting a tie between the two groups. When comparing EM and NM and NF and NM, the results were not significant, and the two groups were tied. Therefore, the hypothesis that their intonation scores were equal cannot be ruled out.

Regarding the rhythm of the interpretations, the EM, NF, and NM groups were 0.333, 0.304, and 0.625 points higher than the mean of the EF group, with a significant difference of at least 5% in all three comparisons. The comparison between EM and NF showed no significant differences, so the hypothesis that these two groups have the same results cannot be ruled out. In the comparison between EM and NM, a significant difference of 0.292 points was found with a confidence level of at least 5% in favor of the NM group. When comparing the NM and NF groups, a significant difference of -0.321 was found, with a confidence level of 5%, in favor of the HN group.

For the interpreters' voice rating, the EM, NF, and NM groups scored on average 0.500, 0.705, and 0.702 points higher than the EF group, respectively, with a significant difference of at least 5% in all three cases. For the first time, when comparing the EM and NF groups,

the NF group scored 0.205 points higher than the EM group, but only at the 10% confidence level. The comparison between EM and NM showed a significant difference of 0.202 points in favor of the beginners, which is a difference of 10%. The last comparison category between the NM and NF groups showed no significant difference. Therefore, the hypothesis that the interpreters' voices of the NF and NM groups were perceived as equally pleasant cannot be rejected.

The eighth parameter in this analysis was perceived interpreter credibility. The means of the EM, NF, and NM groups were 0.269, 0.359, and 0.455 points higher than those of the EF group, respectively, at a confidence level of at least 5% in all comparisons. The last three comparisons of credibility indices were for the EM and NF, EM and NM, and NM vs. NF comparisons, where no significant differences were found. Therefore, the hypothesis that the means of these three groups were equal cannot be ruled out.

Regarding professionalism, there was a noticeable difference of 5% or more between EF and EM, NF, and NM. The mean scores of these three groups were higher than EF by 0.317, 0.359, and 0.474 points. As in the case of credibility, there were no significant differences between EM and NF, EM and NM, or NM and NF. Therefore, we cannot reject the hypothesis that the means of these three groups are equal. For more detailed results of the Tukey's test, please refer to Table 6.

Table 6

Mean difference, standard error, significance, and two-way ANOVA of the study variables.

Independent variable	Difference in means	Std. error	95% confidence interval		Independent variable	Difference in means	Std. error	95% confidence interval		Independent variable	Difference in means	Std. error	95% confidence interval		Sig.	95% confidence interval
			Lower	Upper				Lower	Upper				Lower	Upper		
1. Rate the quality of the interpretation			2. Rate the coherence in the transmission of the interpretation			3. Rate the fluency of the interpretation										
Experienced female	EM	-.349*	.077	.000	-.55	-.15	.077	.000	-.55	-.15	EM	-.218*	.077	.000	-.46	-.04
	NF	-.369*	.077	.000	-.57	-.17	.077	.000	-.57	-.17	NF	-.196	.082	.081	-.41	.02
	NM	-.577*	.077	.000	-.77	-.38	.077	.000	-.77	-.38	NM	-.506*	.082	.000	-.72	-.30
Experienced male	EM	.349*	.077	.000	.15	.55	.078	.027	-.42	-.02	EM	.250*	.082	.013	.04	.46
	NF	-.019	.077	.994	-.22	.18	.078	.000	-.53	-.13	NF	.054	.082	.910	-.16	-.27
	NM	-.228*	.077	.016	-.42	-.03	.078	.000	-.66	-.26	NM	-.256*	.082	.010	-.47	-.05
Novice female	EM	.369*	.077	.000	.17	.57	.078	.000	.17	.42	EM	.196	.082	.081	-.02	.41
	EM	.019	.077	.994	-.18	.22	.078	.448	-.08	.32	EM	-.054	.082	.910	-.27	.16
	NM	.208*	.077	.034	-.41	-.01	.078	.352	-.33	.07	NM	-.311*	.082	.001	-.52	-.10
Novice male	EM	.577*	.077	.000	.38	.77	.078	.000	.26	.66	EM	.506*	.082	.000	.30	.72
	EM	.228*	.077	.016	.03	.42	.082	.010	.04	.44	EM	.256*	.082	.010	.05	.47
	NF	.208*	.077	.034	.01	.41	.078	.352	-.07	.33	NF	.311*	.082	.001	.10	.52

4. Rate the diction of the interpreter										5. Rate the intonation of the interpreter					6. Rate the rhythm maintained by the interpreter				
Experi- enced female	EM	-.151	.080	.231	-.36	-.05	EM	-.375*	.085	.000	-.59	-.16	EM	-.333*	.087	.001	-.56	-.11	
	NF	-.330*	.080	.000	-.53	-.13	NF	-.417*	.085	.000	-.63	-.20	NF	.087	.003	-.53	-.08		
	NM	-.484*	.080	.000	-.69	-.28	NM	-.394*	.085	.000	-.61	-.18	NM	.087	.000	-.85	-.40		
Experi- enced male	EM	.151	.080	.231	-.05	.36	EM	.375*	.085	.000	.16	.59	EM	.333*	.087	.001	.11	.56	
	NF	-.179	.080	.109	-.38	.03	NF	-.042	.085	.961	-.26	.18	NF	.087	.987	-.19	.25		
	NM	-.333*	.080	.000	-.54	-.13	NM	-.019	.085	.996	-.24	.20	NM	.087	.004	-.51	-.07		
Novice female	EM	.330*	.080	.000	.13	.53	EM	.417*	.085	.000	.20	.63	EM	.304*	.087	.003	.08	.53	
	EM	.179	.080	.109	-.03	.38	EM	.042	.085	.961	-.18	.26	EM	.087	.987	-.25	.19		
	NM	-.154	.080	.215	-.36	-.05	NM	.022	.085	.994	-.20	.24	NM	.087	.001	-.54	-.10		
Novice male	EM	.484*	.080	.000	.28	.69	EM	.394*	.085	.000	.18	.61	EM	.625*	.087	.000	.40	.85	
	EM	.333*	.080	.000	.13	.54	EM	.019	.085	.996	-.20	.24	EM	.087	.004	.07	.51		
	NF	.154	.080	.215	-.05	.36	NF	-.022	.085	.994	-.24	.20	NF	.087	.001	.10	.54		

7. Rate the voice of the interpreter										8. Rate the credibility of the interpreter										9. Rate the professionalism of the interpreter									
Ex- peri- enced fe- male	EM	-.500*	.083	.000	-.71	-.29	EM	-.269*	.084	.007	-.49	-.05	EM	-.317*	.080	.000	-.52	-.11											
	NF	-.705*	.083	.000	-.92	-.49	NF	-.359*	.084	.000	-.57	NF	-.359*	.080	.000	-.57	-.15												
	NM	-.702*	.083	.000	-.91	-.49	NM	-.455*	.084	.000	-.67	NM	-.474*	.080	.000	-.68	-.27												
Ex- peri- enced male	EM	.500*	.083	.000	.29	.71	EM	.269*	.084	.007	.05	.49	EM	.317*	.080	.000	.11	.52											
	NF	-.205	.083	.064	-.42	.01	NF	-.090	.084	.708	-.31	NF	-.042	.080	.955	-.25	.17												
	NM	-.202	.083	.070	-.41	.01	NM	-.186	.084	.120	-.40	NM	-.157	.080	.207	-.36	.05												
Nov- ice fe- male	EM	.705*	.083	.000	.49	.92	EM	.359*	.084	.000	.14	.57	EM	.359*	.080	.000	.15	.57											
	EM	.205	.083	.064	-.01	.42	EM	.090	.084	.708	-.13	EM	.042	.080	.955	-.17	.25												
	NM	.003	.083	1.000	-.21	.22	NM	-.096	.084	.661	-.31	NM	-.115	.080	.479	-.32	.09												
Nov- ice male	EM	.702*	.083	.000	.49	.91	EM	.455*	.084	.000	.24	.67	EM	.474*	.080	.000	.27	.68											
	EM	.202	.083	.070	-.01	.41	EM	.186	.084	.120	-.03	EM	.157	.080	.207	-.05	.36												
	NF	-.003	.083	1.000	-.22	.21	NF	.096	.084	.661	-.12	NF	.115	.080	.479	-.09	.32												

Note: Sig. = significance: n = 156.
 * The mean difference was significant, with a confidence level of 0.05 %.

4. DISCUSSION AND CONCLUSIONS

The age and sex of the interpreters were clearly distinguished in the study variables, as shown in the explanation of the results above, from which the following conclusions can be drawn:

1) Novice males consistently ranked first in the nine categories, and experienced females ranked last. Significant differences were found between the EF group and the other groups. The EM, NF, and NM groups were always favored. These results corroborate the Baseline Research. This marked differentiation can be explained by the fact that anatomical variations related to the age and sex of the speaker produce significant changes in the fundamental frequency (F0) and formant frequencies (Baus et al., 2019; Sundberg, 2019). These two acoustic features, which are the most important for identifying the sex and age of an individual, are recorded in less than 140 ms from the onset of the utterance (Schweinberger et al., 2008). The sex of the speaker is readily recognized and encoded in brain areas specialized in the perception of male or female voices before the way words are articulated or the linguistic and grammatical components of the speaker are recognized (Greenberg & Christiansen, 2019). Voice characteristics influence how listeners perceive the speaker and respond to the information they convey. This process can influence the perception of an SI's verbal and nonverbal characteristics, which was unfavorable for the mature women in this study.

2) The male novice interpreters outperformed all their peers in three categories: quality, rhythm, and fluency. None of the participants in the other three groups significantly outperformed their peers. The perception of performance quality is related to several parameters (Iglesias Fernández, 2010), such as the interpreter's voice, fluency, and rhythm (Kiktová et al., 2019) of the rendition, where in this study, novice males scored the highest. This result corroborates the findings of the Baseline Research, where novice interpreters outperformed their peers in terms of quality, rhythm, professionalism, credibility, and voice. A higher rating for several parameters in the same age group can be interpreted as a preference of the listeners for this group.

3) No experienced interpreters, male or female, obtained significantly higher means than the novice interpreters.

4) Novice female interpreters were comparable to experienced male interpreters in all parameters except voice. Lower voices, generally male, score higher due to their ability to evoke greater credibility and persuasiveness, especially when conveying descriptive and informative messages (Larrea Estefanía, 2014), as in the case of the SI conducted in this study. These results show that there can be interactions between voice and sexual stereotypes (Strand, 2000). As mentioned above, most people integrate social stereotypes into their perception of speech, which elicits automatic, almost instinctive attitudes and judgments toward speakers (Strand, 2000).

5) Experienced male interpreters were comparable to female and male novices in intonation, professionalism, and credibility, yet all three groups outperformed experienced female interpreters. This perception of more excellent professionalism of male experienced interpreters than female experienced interpreters may be due to the content and activation of sexual stereotypes (Casper & Rothermund, 2012). When investigating age-related changes, Kornadt *et al.* (2013) showed that in most domains of experience, these changes were more favorable for women than men (family, friends, religion, leisure, personality, and health). However, in the work and finance domains, a reverse effect was identified that favored men over women. This double standard may also be due to stigmatizing attitudes towards women (Barret & Naiman-Sessions, 2016; Krekula *et al.*, 2018), highlighting the dominance of patriarchal norms combined with a more pronounced concern about age in older women.

6) When the mean scores were disaggregated by group, the same results were found, with beginners systematically ranked first and experienced interpreters ranked last in all parameters. The means of G2 were systematically higher than those of G1 (the disaggregated analysis is not included here as it is beyond the scope of this publication).

The results of this research corroborate the conclusions of the Baseline Research. The following empirical conclusions can be drawn from the findings with statistically significant differences in scores according to the age and sex of the interpreter: 1) male interpreters tend

to score higher than their female counterparts of the same age; 2) novice interpreters tend to score higher than experienced interpreters, regardless of sex; and 3) comparisons between age groups are less favorable for experienced interpreters.

Native speakers of a given language acquire a specific awareness of how phonological categories are uttered, depending not only on the different communities that use it but also on the sex of the utterer (Madrid Servín & Marín Rodríguez, 2001). Thus, it cannot be ruled out that the peculiarities or culture of the country where the study was conducted may encourage sexist or ageist stereotypes, especially towards mature women. Further research is therefore needed beyond the country where the study was conducted to determine whether such ageist stereotypes are prevalent elsewhere.

Experience, an indisputable resource for all professional interpreters (Moser-Mercer, 2022), is undoubtedly relevant to their performance. However, the findings suggest that when two interpreters sharing the same booth, have the same level of preparedness and similar knowledge of the interpreted topic, it is likely that, if there are any presbyphonic elements in the voice of either of them, differences in the perception of their interpretations will be found. The statements expressed here can be applied when listeners only have the interpreter's voice as a reference rather than other professional skills that could be perceived as an added value to the service.

The empirical results of this study do not allow us to determine whether the sex or age of the interpreters, by themselves, are determining factors in the perception of SI, i.e., whether one has more weight than the other or whether they are due to a reciprocal effect depending on the age group. However, these data suggest that before the verbal elements of SI are perceived, the interaction between the perceived sex and age of the interpreter unconsciously biases the listener's evaluation of the interpretation. This bias may be due to the listener's cultural norms and stereotypes based on the vocal characteristics of the interpreter.

Declaration of Conflict of Interest

The author declares that she has no potential conflict of interest concerning the research, authorship, and publication of this article.

Funding

This research did not receive any specific grant from any funding body in the public, commercial, or non-profit sectors.

4. ACKNOWLEDGMENTS

Special thanks to Dr. Almudena Nevado Llopis of the Universidad de San Jorge for her guidance in reviewing this article.

REFERENCES

- Abdulsatar, A.A., Davydov, V.V., Yushkova, V.V., Glinushkin, A.P., & Rud, V.Y. (2019). Age and gender recognition from speech signals. *Journal of Physics: Conference Series*, 1410(1), 0-7. <https://doi.org/10.1088/1742-6596/1410/1/012073>
- Angelelli, C.V. (2004). *Revisiting the Interpreter's Role*. John Benjamins Publishing Company. <https://doi.org/10.1075/btl.55>
- Barret, A., & Naiman-Sessions, M. (2016). 'It's our turn to play': Performance of girlhood as a collective response to gendered ageism. *Aging and Society*, 36(4), 764-784. <https://doi.org/10.1017/s0144686x15000021>
- Bartłomiejczyk, M. (2020). Parliamentary impoliteness and the interpreter's gender. *Pragmatics*, 30(4), 459-484. <https://doi.org/10.1075/prag.18064.bar>
- Baus, C., McAleer, P., Marcoux, K., Belin, P., & Costa, A. (2019). Forming social impressions from voices in native and foreign languages. *Scientific Reports*, 9(1), 1-14. <https://doi.org/10.1038/s41598-018-36518-6>

- Bhargava, A., Arnold, A.P., Bangasser, D.A., Denton, K.M., Gupta, A., Hilliard Krause, L.M., Mayer, E.A., McCarthy, M., Miller, W.L., Raznahan, A., & Verma, R. (2021). Considering sex as a biological variable in basic and clinical studies: An endocrine society scientific statement. *Endocrine Reviews*, 42(3), 219-258. <https://doi.org/10.1210/endrev/bnaa034>
- Blasco Mayor, M.J., & García Becerra, O. (2007). La Incidencia del Parámetro Dicción. In A. Collados Aís, M. Pradas Macías, E. Stévaux, & O. García Becerra (Eds.), *La Evaluación de la Calidad en Interpretación Simultánea: parámetros de incidencia* (pp. 175-194). Comares.
- Casper, C., & Rothermund, K. (2012). Gender self-stereotyping is context dependent for men but not for women. *Basic Applied Social Psychology*, 34(5), 434-442. <https://doi.org/10.1080/01973533.2012.712014>
- Cheung, A.K.F. (2020). Interpreters' perceived characteristics and perception of quality in interpreting. *Interpreting*, 22(1), 35-55. <https://doi.org/10.1075/intp.00033.che>
- Chmiel, A. (2021). Effects of simultaneous interpreting experience and training on anticipation, as measured by word-translation latencies. *Interpreting. International Journal of Research and Practice in Interpreting*, 23(1), 18-44. <https://doi.org/10.1075/intp.00048.chm>
- Christen y Gracia, L. (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* (pp. 26-58). Comares.
- Collados Aís, A. (2019). La entonación monótona y la calidad de la interpretación simultánea: frecuencia, conceptualizaciones y efectos. *Meta*, LXI(3), 675-691. <https://doi.org/10.7202/1039224ar>
- Collados Aís, A., Pradas Macías, M., Stévaux, E., & García Becerra, O. (2007). *La Evaluación de la Calidad en Interpretación Simultánea: Parámetros de Incidencia*. Comares.
- Collard, C., & Defrancq, B. (2017). *Sex Differences in Simultaneous Interpreting: A Corpus-Based Study*. UGent. <http://hdl.handle.net/1854/LU-8518872>
- Collard, C., & Defrancq, B. (2019). Predictors of ear-voice span, a corpus-based study with special reference to sex. *Perspectives*, 27(3), 431-454. <https://doi.org/10.1080/0907676X.2018.1553199>
- Cuddy, A.J.C., & Fiske, S.T. (2004). Doddering but dear: Process, content, and function in stereotyping of older persons. In T.D. Nelson (Ed.), *Ageism. Stereotyping and Prejudice against Older Persons* (pp. 1-26). The MIT Press.
- Defrancq, B., Collard, C., Magnifico, C., & Iglesias Fernández, E. (2021). Sex and gender in conference interpreting. En M. Albl-Mikasa & E. Tiselius (Eds.), *The Routledge Handbook of Conference Interpreting* (pp. 414-427). The Routledge Taylor and Francis Group.

- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford University Press.
- García Becerra, O. (2012). *La incidencia de las primeras impresiones en la evaluación de la calidad de la interpretación simultánea: un estudio empírico*. Editorial de la Universidad de Granada. <https://digibug.ugr.es/bitstream/handle/10481/25153/21620234.pdf?sequence=1>
- García Becerra, O. (2015). Order Effect, Impression Formation, and Their Impact on the Evaluation of Interpreting Quality. In C. Zwischenberger & M. Behr (Eds.), *Interpreting Quality: A Look Around and Ahead* (pp. 123-146). Frank & Timme.
- Gentile, P. (2016). The Interpreter's Professional Status. A Sociological Investigation into the Interpreting Profession. [Unpublished doctoral thesis]. Università degli studi di Trieste. <https://hdl.handle.net/11368/2908044>
- Gile, D. (2018). *THE CIRIN BULLETIN. Conference Interpreting Research Information Network*. Bulletin no. 56. <http://cirinandgile.com/bulletins/Bulletin-56-Jul-2018.pdf>
- Gile, D. (2002). Conference Interpreting as a Cognitive Management Problem. In F. Pöchhacker & M. Shlesinger (Eds.), *The Interpreting Studies Reader* (pp. 163-176). Routledge Taylor & Francis Group.
- Greenberg, S., & Christiansen, T. (2019). The perceptual flow of phonetic information. *Attention, Perception, and Psychophysics*, 81, 884-896. <https://doi.org/10.3758/s13414-019-01666-y>
- Horváth, I. (2017). The speech behaviour of interpreters. *Across Languages and Cultures*, 18(2), 219-236. <https://doi.org/10.1556/084.2017.18.2.3>
- Hummert, M., Mazloff, D., & Henry, C. (1999). Vocal characteristics of older adults and stereotyping. *Journal of Nonverbal Behavior*, 23(2), 111-132. <https://doi.org/10.1023/A:1021483409296>
- Iglesias Fernández, E. (2007). La Incidencia del Parámetro Agradabilidad de la Voz. In A. Collados Aís, M. Pradas Macías, E. Stévaux, & O. García Becerra (Eds.), *La Evaluación de la Calidad en Interpretación Simultánea: parámetros de incidencia* (pp. 37-51). Comares.
- Iglesias Fernández, E. (2010). Understanding variability in interpreting quality assessment: User's sex and judgments for pleasant voice. In C. Way, S. Vandepitte, R. Meilaerts, & M. Bartolomiejczyk (Eds.), *Tracks and Treks in Translation Studies: Selected papers from the EST Congress, Leuven* (pp. 103-125). John Benjamins.
- Kiktová, E., Zimmermann, J., & Palová, M. (2019). Towards to the Anticipation in Simultaneous Interpreting. In *CEUR Workshop Proceedings* (Vol. 2473, pp. 171-175). CEUR-WS. <https://ceur-ws.org/Vol-2473/paper35.pdf>

- Kornadt, A.E., Voss, P., & Rothermund, K. (2013). Multiple standards of aging: Gender-specific age stereotypes in different life domains. *European Journal of Ageing*, 10(4), 335-344. <https://doi.org/10.1007/s10433-013-0281-9>
- Krekula, C., Nikander, P., & Wilińska, M. (2018). Multiple Marginalizations Based on Age: Gendered Ageism and Beyond. In L. Ayalon & C. Tesch-Römer (Eds.), *Contemporary Perspectives on Ageism* (pp. 33-51). Springer Nature.
- Larrea Estefanía, O. (2014). *Estudio sobre la escucha de la voz del locutor con y sin su imagen: análisis del proceso perceptivo y cognitivo del mensaje*. [Unpublished doctoral thesis]. Universidad Pompeu Fabra. <http://hdl.handle.net/10803/292736>
- Liu, M. (2008). How do experts interpret? Implications from research in Interpreting Studies and cognitive science. In G. Hansen, A. Chesterman, & H. Gerzymisch-Arbogast (Eds.), *Efforts and Models in Interpreting and Translation Research* (pp. 159-177). John Benjamins Publishing Company.
- Liu, M., Kurz, I., Moser-Mercer, B., & Shlesinger, M. (2020). The interpreter's aging: A unique story of multilingual cognitive decline? *Translation, Cognition and Behavior*, 3(2), 287-310. <https://doi.org/10.1075/tcb.00045.liu>
- Madrid Servín, E., & Marín Rodríguez, M. (2001). Estructura formántica de las vocales del español de la Ciudad de México. In E. Herrera (Ed.), *Temas de Fonética Instrumental* (pp. 39-53). El Colegio de México.
- Magnifico, C. (2017). Hedges in Conference Interpreting: The role of gender. *Interpreting*, 191, 21-46. <https://doi.org/10.1075/intp.19.1.02mag>
- Magnifico, C., & Defrancq, B. (2020). Norms and gender in simultaneous interpreting: A study of connective markers. *Translation and Interpreting*, 12(1), 1-17. <https://doi.org/10.12807/TI.112201.2020.A01>
- Moser-Mercer, B. (2022). Conference interpreting and expertise. In M. Albl-Mikasa & E. Tiselius (Eds.), *The Routledge Handbook of Conference Interpreting* (pp. 386-400). Routledge Taylor and Francis Group.
- Nakane, I. (2008). Politeness and gender in interpreted police interviews. *Monash University Linguistics Papers*, 6(1), 29-40.
- Nass, C., & Brave, S. (2005). *Wired for Speech. How Voice Activates and Advances the Human-Computer Relationship*. The MIT Press.
- Palová, M., & Kiktová, E. (2019). Prosodic anticipatory clues and reference activation in simultaneous interpretation. *XLinguae*, 12(IXL), 13-22. <https://doi.org/10.18355/XL.2019.12.01XL.02>

- Pisanski, K., & Feinberg, D. (2015). Vocal Attractiveness. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (pp. 607-625). Oxford University Press.
- Pöschhacker, F. (2016). *Introducing Interpreting Studies*. Routledge.
- Rennert, S. (2020). Impact of fluency from a listener perspective. In R. Barranco-Droege (Ed.), *Solving the riddle of interpreting quality*. Comares.
- Riccardi, A. (2022). Strategies and capacity management in conference interpreting. In M. Albl-Mikasa & E. Tiselius (Eds.), *The Routledge Handbook of Conference Interpreting* (pp. 371-385). Routledge Taylor and Francis Group.
- Schweinberger, S.R., Casper, C., Hauthal, N., Kaufmann, J.M., Kawahara, H., Kloth, N., Robertson, D.M.C., Simpson, A.P., & Zäske, R. (2008). Auditory Adaptation in Voice Perception. *Current Biology*, 18, 684-688. <https://doi.org/10.1016/j.cub.2008.04.015>
- Strand, E. (2000). *Gender stereotype effects on speech processing*. (Publicación No. 9994943) [Doctoral thesis, Ohio State University]. <https://www.proquest.com/openview/03f3eea67ec3a4b78ec5b95c94e8b19a/1?pq-origsite=gscholar&cbl=18750&diss=y>
- Streiner, D.L. (2015). Best (but oft-forgotten) practices: The multiple problems of multiplicity-whether and how to correct for many statistical tests. *American Journal of Clinical Nutrition*, 102(4), 721-728. <https://doi.org/10.3945/ajcn.115.113548>
- Sundberg, J. (2019). The Singing Voice. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (pp. 117-142). Oxford University Press.
- Tiselius, E. (2013). *Experience and Expertise in Conference Interpreting. An Investigation of Swedish Conference Interpreters*. [Unpublished doctoral thesis]. University of Bergen. <https://bora.uib.no/bora-xmlui/bitstream/handle/1956/8747/dr-thesis-2013-Elisabet-Tiselius.pdf?sequence=1>
- Varela García, M., Ruiz Rosendo, L., & Barghout, A. (2019). Speed in simultaneous interpreting. In *100 years of conference interpreting*. Ginebra. <https://archive-ouverte.unige.ch/unige:129083>
- Voss, P., Bodner, E., & Rothermund, K. (2018). Ageism: The Relationship between Age Stereotypes and Age Discrimination. In L. Ayalon & C. Tesch-Römer (Eds.), *Contemporary Perspectives on Ageism* (pp. 11-31). Springer, Cham. https://doi.org/10.1007/978-3-319-73820-8_2
- Zhang, H., Liu, M., Li, W., & Sommer, W. (2020). Human voice attractiveness processing: Electrophysiological evidence. *Biological Psychology*, 150, 107827. <https://doi.org/10.1016/j.biopsycho.2019.107827>