

Article

Gender and Accent Biases in AI-Based Tools for Spanish: A Comparative Study between Alexa and Whisper

Eduardo Nacimiento-García *, Holi Sunya Díaz-Kaas-Nielsen and Carina S. González-González *

Women Studies Research Institute (IUEM), University of La Laguna, 38200 La Laguna, Spain; hdiakaa@ull.edu.es

* Correspondence: enacimie@ull.edu.es (E.N.-G.); cjgonza@ull.edu.es (C.S.G.-G.)

Abstract: Considering previous research indicating the presence of biases based on gender and accent in AI-based tools such as virtual assistants or automatic speech recognition (ASR) systems, this paper examines these potential biases in both Alexa and Whisper for the major Spanish accent groups. The Mozilla Common Voice dataset is employed for testing, and after evaluating tens of thousands of audio fragments, descriptive statistics are calculated. After analyzing the data disaggregated by gender and accent, it is observed that, for this dataset, in terms of means and medians, Alexa performs slightly better for female voices than for male voices, while the opposite is true for Whisper. However, these differences in both cases are not considered significant. In the case of accents, a higher Word Error Rate (WER) is observed among certain accents, suggesting bias based on the spoken Spanish accent.

Keywords: accent bias; Alexa; ASR; automatic speech recognition; gender bias; Whisper

1. Introduction

In recent years, we have witnessed a surge in the utilization of speech-recognition technologies and voice interaction [1,2]. One of the domains where voice interaction is being employed is that of virtual assistants. Prominent commercial voice assistants include Amazon Alexa, Google Assistant, and Apple's Siri [3], with Alexa being the most prevalent, occupying approximately 70% of the market share [4].

In addition to the aforementioned commercial virtual assistants and open-source software like Home Assistant [5], a variety of automatic speech recognition (ASR) tools are available, enabling the implementation of a system that transcribes speech to text (STT) [6]. This capability can be harnessed for the development of voice interaction-based systems, such as virtual assistants. Notably, among the speech-recognition systems, Whisper [7] stands out, as its introduction has prompted similar open-source tools like Coqui STT [8] to discontinue their projects due to the improvements offered by this new tool. Whisper is free software.

From the perspective of human–device interaction through speech, it is crucial to consider several key concepts that set it apart from other forms of interaction [9]. However, in our context, we will focus on the importance of the device accurately understanding the individual in their language, dialectal variation, and accent. Furthermore, it is essential to ensure that there is no significant difference in performance when these devices are used by both females and males.

Currently, it is estimated that approximately 8.1 billion people inhabit the world [10]. Nevertheless, there is no single language that is spoken or understood by the entire global population, not even by a majority. According to data published by Ethnologue [11], the most widely spoken language is English, considering both native speakers and those who speak it as a second language. English is spoken by approximately 1.456 billion people, which roughly equates to 18% of the world's population. If we consider the top 10 most spoken languages (English, Mandarin Chinese, Hindi, Spanish, Standard Arabic, Bengali,



Citation: Nacimiento-García, E.; Díaz-Kaas-Nielsen, H.S.; González-González, C.S. Gender and Accent Biases in AI-Based Tools for Spanish: A Comparative Study between Alexa and Whisper. *Appl. Sci.* **2024**, *14*, 4734. <https://doi.org/10.3390/app14114734>

Academic Editor: Andrea Prati

Received: 21 February 2024

Revised: 23 May 2024

Accepted: 28 May 2024

Published: 30 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

French, Russian, Portuguese, and Urdu), we will encompass approximately 66% of the global population, leaving over 2.7 billion people (34%). The top 200 languages spoken account for approximately 88% of the world's population [8].

Access to information and human knowledge by all individuals, regardless of their language, is paramount and should be regarded as a fundamental right. Thanks to the Internet, access to a portion of information and human knowledge has become more democratic, yet much remains to be accomplished [12].

According to UNESCO, roughly 781 million people worldwide are illiterate, with approximately two thirds of them being women [13]. Hence, voice interaction technology must be adept enough to interpret a wide range of languages, dialectal variations, and accents, regardless of gender.

Previous studies have revealed historical automatic speech recognition (ASR) system biases [14,15]. These biases hinder effective communication for certain groups of people when using voice recognition systems [16]. Some of these biases [17,18] may be attributed to cultural, social, medical, or other differences, making gender and dialectal variations or accents of the interacting individuals two significant sources of potential bias within ASR systems [19,20].

In this context, it is relevant to consider potential gender biases [21,22] in speech-recognition tools and virtual assistants, both in the responses provided by an assistant and in the actual voice recognition. This study will specifically focus on speech recognition.

In general terms, gender identification through voice primarily relies on fundamental frequency [23]. On average, female voices have a fundamental frequency of approximately one octave higher than male voices [24]. Fundamental frequency refers to the lowest vibration frequency of the vocal cords during sound production. Typically, female vocal cords tend to be shorter and thinner than male vocal cords, resulting in a higher fundamental frequency in female voices and a lower one in male voices [25].

As previously mentioned, in addition to the various languages, different dialects exist within languages. In our case, we will focus on the dialects of the Spanish language. One of the current classifications identifies eight dialectal regions of the Spanish language [26], with five in the Americas, two in Europe, and one in Africa. The dialectal regions encompass the following areas: America, which includes the Caribbean, Mexican-Central American, Andean, Austral, and Chilean regions; Europe, consisting of the Northern Iberian Peninsula (Septentrional) and the Southern Iberian Peninsula (Meridional); and Africa, which comprises the Canary Islands.

This research aims to ascertain whether there is any significant bias concerning gender or the main accents of the Spanish language. To achieve this, audio clips from the Common Voice 14 dataset by Mozilla [27] are analyzed using both Alexa and Whisper.

Section 3 introduces the tools and datasets employed for the analyses. Section 4 presents the outcomes of the tests conducted. In Section 5, a discussion of the obtained results is provided. Finally, Section 6 presents the conclusions drawn from this research.

2. Background

The examination of various technologies and tools based on Artificial Intelligence (AI) has revealed that, in many instances, different types of biases exist or have existed, adversely affecting specific social groups compared to others. For instance, biases have been identified concerning membership in various ethnic groups, the use of different accents or dialectal variants, and gender, among other factors [28–30]

Specifically, biases have also been detected in automatic speech recognition (ASR), both based on the speaker's gender and the accent or dialectal variant used by individuals speaking the same language [31].

Often, these biases stem from the data itself used to train AI-based systems. Consequently, bias resulting from such data could be readily mitigated by employing datasets that are not already predisposed to bias [32]. Studies related to the topic include comparisons between English accents, such as American and Indian, along with considerations of

gender. In an evaluation of the DeepSpeech (STT) tool, bias was found based on accent, although no gender bias was observed in this case [33]. In another study, the transcription performed by YouTube using ASR for different English accents was examined, revealing biases in both gender and accent [31]. Similarly, a study demonstrated the existence of gender bias unfavorable to women in some ASR systems, attributed to the use of biased data in model training.

This resulted in a higher Word Error Rate (WER) for females when interacting with these systems compared to males [34]. These prior investigations confirm the need to continue research in this field to ascertain whether such biases persist in these ASR tools or if, conversely, there has been an evolution with a reduction or elimination of such biases.

3. Materials and Methods

3.1. Objectives

Building upon the considerations outlined in Section 2 (Background), the primary objective of this research is to verify the existence of gender or accent bias in automatic speech recognition (ASR) for the Spanish language when utilizing the voice-activated virtual assistant Alexa or the Whisper system.

The aim is to gather data that facilitates comparisons based on gender and the primary accents or dialectal variations in Spanish. Through this investigation, we seek to discern potential biases in the ASR systems and contribute insights into how these biases may vary concerning gender and diverse linguistic features within the Spanish language.

3.2. Inclusion and Exclusion Criteria

The analysis was conducted on many audio segments sourced from the Mozilla Common Voice dataset for the Spanish language, version 14, released on 28 June 2023 [27]. This dataset comprises 1,608,353 audio fragments and their corresponding text transcriptions, making it suitable for applications such as automatic speech recognition. The dataset encompasses 2175 recorded hours, of which 504 h are validated, involving the participation of 25,261 contributors. In this research, we employed the Common Voice 14 dataset in Spanish to investigate speech recognition in the Spanish language, particularly in identifying potential biases related to gender and accent usage.

The Common Voice dataset is subject to human validation, where a segment is considered correct when it receives two positive votes and incorrect when it garners two negative votes. There is also a possibility that some segments may receive both positive and negative evaluations simultaneously. During the data refinement process for the Common Voices dataset, we searched to select segments meeting the following criteria:

- Must have at least two positive votes and no negative votes to exclude uncertain segments.
- Segments are categorized by gender: female and male; this categorization is optional, and not all segments are labeled by gender.
- Segments are categorized based on the speaker's accent, as this is also an optional feature when contributing to the Mozilla Common Voice project.

Once we obtained the dataset consisting of segments categorized by gender and accent and validated it without negative votes, we analyzed the various accents available.

In this case, it should be noted that not all segments use a standardized category for accent classification. This is likely due to the initial stage of the project when the accent field may have been an open-text field rather than a selection field as it currently is. Another filter applied was to ensure that the audio segments consisted of more than one word, as there are segments with single words such as "Yes", "No", or numbers, which could significantly influence the final analysis results as an error in a single word translates to an entire sentence error.

To obtain a sample with a manageable dataset, we proceeded to filter and retain categories where there were at least 500 audio segments for each accent and gender.

Following this final filtering, the following categories based on accent were retained:

- Central American
- Andean-Pacific: Colombia, Peru, Ecuador, Western Bolivia, Andean Venezuela
- Caribbean: Cuba, Venezuela, Puerto Rico, Dominican Republic, Panama, Caribbean Colombia, Caribbean Mexico, Gulf Coast of Mexico
- Chilean: Chile, Cuyo
- Northern Iberian Peninsula (Asturias, Castilla y León, Cantabria, Basque Country, Aragon, La Rioja, Guadalajara, Cuenca)
- Central-Southern Iberian Peninsula (Madrid, Toledo, Castilla-La Mancha)
- Southern Iberian Peninsula (Andalusia, Extremadura, Murcia)
- Canary Islands
- Mexico
- Rioplatense: Argentina, Uruguay, Eastern Bolivia, Paraguay

After filtering, the total dataset consisted of 202,737 audio segments, which were distributed as follows, as shown in Table 1. Many segments categorized with non-standardized labels were left out, and the only current category available in the Common Voice selection is “Español de Filipinas”. In the dataset, there are only 23 segments for this accent, with 10 from males and 13 from females, and after filtering, it would be reduced to 10 and 5, respectively.

Table 1. Number of segments by gender and accent.

Accent	Female	Male
Central American	1501	3908
Andean-Pacific	2016	11,493
Caribbean	1537	6041
Chilean	1268	3939
Spain Northern	3610	31,406
Spain Central	1958	6930
Spain Southern	1456	30,698
Canary Islands	562	1482
Mexico	34,639	47,020
Rioplatense	2914	8359

The final accent-based categories we have retained closely align with the previously referenced dialectal classification [26], with the caveat that, in this context, some accents are further subdivided. The updated classification is as follows: The “Andean” accent corresponds to “Andean”, while the “Canary” accent is linked to the “Canary Islands”. The “Caribbean” accent aligns with “Caribbean”, and the “Chilean” accent is associated with “Chileno”. The “Meridional Spanish” accent is attributed to both the “Southern Iberian Peninsula” and the “Central-Southern Iberian Peninsula” categories. “Septentrional Spanish” corresponds to the “Northern Iberian Peninsula”, and “Mexico-Central America” is divided into the “Central America” and “Mexico” categories. Lastly, the “Austral” accent is affiliated with the “Rioplatense” category.

3.3. Tools

As highlighted in Section 1, Alexa stands out as the most widely utilized voice-activated virtual assistant [4], and Whisper, in its brief existence, has revolutionized the sector to the extent of influencing established projects such as Coqui STT, leading them to discontinue their development [8]. Currently, Alexa can be regarded as the foremost benchmark for Voice-Activated Virtual Assistants, while Whisper is a crucial reference for automatic speech-recognition (ASR) systems. This underscores the rationale for subjecting the automatic speech recognition of the voice-activated virtual assistant Alexa and the voice-to-text system Whisper to a comprehensive analysis.

Alexa [35] is a voice-activated virtual assistant developed by Amazon, available on specific devices such as Amazon Echo. It is also accessible as a mobile application for both

Android and iOS. Additionally, web access to a developed Skill (App) is possible through the developer console. Alexa is proprietary software.

Whisper [7], on the other hand, is an automatic speech-recognition system developed by OpenAI, licensed under the MIT license, and characterized as open-source software. According to the developers, Whisper has been trained on 680,000 h of multilingual and multitasking supervised data collected from the web [36]. The Base and Large-v2 models of Whisper were chosen from among the six available at the time of the research. The selection of the Large-v2 model was based on its comprehensive nature, akin to the Large model, with the added advantage of being the latest version. Conversely, despite the smallest Tiny model, the decision was made to opt for the Base model, as its name implies, positioning it as the reference model.

In addition to these two main tools, Selenium [37] was employed during the research to automate the process in Alexa. Coqui TTS [38] generated the activation word preceding each phrase in Alexa. Finally, Python 3.10 was the programming language utilized to create various scripts for data analysis in both Alexa and Whisper.

3.4. Procedure

The data analysis procedure is straightforward, primarily involving processing an audio snippet from Common Voice using either Alexa or Whisper. Subsequently, the transcription of the audio is obtained, and this text is then compared with the actual transcription provided by Common Voice. This comparison is achieved by calculating the WER. Finally, after collecting the data, statistical calculations are performed to observe the data jointly based on model variants, gender, or accent variations.

An important point to consider is that Alexa has three Spanish variants: Mexican, US, and Spain Spanish. This requires conducting all analyses in triplicate to ensure a comprehensive evaluation. In the case of Whisper, there is only a generic model for Spanish, although there are different models of various sizes.

A Python script with Selenium [37] was created to analyze Alexa, enabling access to the Amazon Alexa developer console and interacting with the system from there. An external USB sound card was used for hardware to eliminate additional noise during sound capture. A simple Alexa Skill (application) was developed, which received an audio input and processed it to respond. Using the Alexa developer console allowed for straightforward capturing of the resulting text. It is worth noting that when developing the Alexa Skill, the generic data type AMAZON.SearchQuery [39] was used because an open and broad dataset was employed. In this case, using more specific data types could potentially influence the speech-recognition analysis results. The issue that needed to be addressed was that with AMAZON.SearchQuery data type, a triggering word needed to precede the phrase for Alexa to detect it. In this project, Coqui TTS [38] was used to generate an audio segment with the word “Escucha” (Listen) in Spanish, which preceded the segments during the tests. Later on, the word “Escucha” was removed from the obtained texts to avoid impacting the results. One of the project requirements for Common Voice is that the segments consist of fewer than 15 words [27]. After conducting various tests by preceding different phrases with the word ‘Listen,’ no interference was observed. The word ‘Listen’ was correctly detected in all tests.

For Whisper, the process was more straightforward, as the system is designed to be used via a Python program that takes an audio file for transcription. Currently, there are six models available for Whisper [7]. For the analysis, the Base and Large-v2 models of Whisper were used. The available models encompass a range of options, including Tiny, Base, Small, Medium, Large, and Large-v2.

For the analysis, 56,344 segments were processed for Alexa, 202,737 for Whisper (Base), and 105,375 for Whisper (Large-v2). The difference in sample sizes is due to the time constraints of conducting the tests. The Alexa tests took approximately four weeks to complete, while the Whisper Base tests were finished in only four days, and the Whisper Large-v2 tests

took approximately one week. It is worth noting that according to data provided by Whisper’s creators, the Base model is 16 times faster than the Large or Large-v2 model [7].

The samples used are listed in Tables 2 and 3 for females and males, respectively. As shown in the tables, a sample of up to a maximum of 2150 audio segments was used in the case of Alexa. If the number is lower, it is because there were no more segments with that accent for that gender that met the established requirements. For Whisper (Base), the samples consisted of all audio segments that satisfied the specified requirements. In the case of Whisper (Large-v2), the limit was set at 11,600 audio segments. In the cases of Alexa and Whisper Large-v2, the selected segments up to the cutoff limit were chosen randomly.

Table 2. Sample sizes used for females.

Accent	Alexa	Whisper Base	Whisper Large-v2
Andean	2016	2016	2016
Canary	562	562	562
Caribbean	1537	1537	1537
Central American	1501	1501	1501
Chilean	1268	1268	1268
Spain Central	1958	1958	1958
Spain Northern	2150	3610	3610
Spain Southern	1456	1456	1456
Mexican	2150	34,639	11,600
Rioplátense	2150	2914	2915
Total	16,748	51,461	28,423

Table 3. Sample sizes used for males.

Accent	Alexa	Whisper Base	Whisper Large-v2
Andean	2150	11,493	11,493
Canary	1482	1482	1482
Caribbean	2150	6041	6041
Central American	2150	3908	3908
Chilean	2150	3939	3939
Spain Central	2150	6930	6930
Spain Northern	2150	31,406	11,600
Spain Southern	2150	30,698	11,600
Mexican	2150	47,020	11,600
Rioplátense	2150	8359	8359
Total	20,832	151,276	76,952

4. Results

After analyzing the various samples to calculate error rates related to automatic speech-recognition analysis, such as the WER [40], we proceed to compute various statistics, such as the mean, median, standard deviation, and variance. Additionally, we calculated the 95% confidence intervals to ensure that the results obtained were conclusive.

A preprocessing step was performed on the character strings to obtain a WER that is as realistic as possible. This involved removing punctuation marks, exclamation marks, question marks, etc., and converting all strings to lowercase. This ensured that the WER calculation would not be affected by different interpretations between Alexa and Whisper beyond the simple transcription of the words heard.

Tables 4–9 display the results obtained for Alexa, disaggregated by gender and Alexa variant. Tables 10–13 present results disaggregated by gender for Whisper.

Table 4. WER (%): Alexa, MX, Female.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	39.83	28.57	36.02	12.98	1.57
Canary	35.46	25.00	34.71	12.05	2.88
Caribbean	42.43	33.33	37.01	13.70	1.85
Central American	42.45	30.77	36.81	13.55	1.86
Chilean	42.12	33.33	36.35	13.21	2.00
Spain Central	35.76	22.22	36.21	13.11	1.60
Spain Northern	34.32	22.22	35.56	12.65	1.50
Spain Southern	42.90	33.33	35.96	12.93	1.85
Mexican	38.29	25.00	36.38	13.24	1.54
Rioplátense	40.51	28.57	36.66	13.44	1.55

Table 5. WER (%): Alexa, MX, Male.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	43.37	30.77	37.90	14.37	1.60
Canary	41.39	30.00	34.78	12.10	1.77
Caribbean	41.53	27.27	38.82	15.07	1.64
Central American	43.02	30.77	37.33	13.93	1.58
Chilean	40.47	27.27	37.45	14.03	1.58
Spain Central	41.94	28.57	38.53	14.84	1.63
Spain Northern	41.37	27.92	38.71	14.99	1.64
Spain Southern	47.25	35.71	39.53	15.62	1.67
Mexican	44.71	30.77	39.33	15.47	1.66
Rioplátense	40.01	27.27	37.72	14.23	1.60

Table 6. WER (%): Alexa, ES, Female.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	36.72	25.00	33.38	11.14	1.46
Canary	42.52	30.00	37.77	14.27	3.13
Caribbean	40.58	30.77	35.51	12.61	1.78
Central American	39.64	28.57	34.46	11.88	1.74
Chilean	41.15	30.77	35.60	12.67	1.96
Spain Central	33.42	22.22	33.54	11.25	1.49
Spain Northern	31.70	22.22	31.95	10.21	1.35
Spain Southern	40.50	33.33	33.32	11.11	1.71
Mexican	34.07	23.08	33.42	11.17	1.41
Rioplátense	37.45	28.57	33.80	11.43	1.43

Table 7. WER (%): Alexa, ES, Male.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	40.06	30.00	35.47	12.58	1.50
Canary	40.82	30.77	32.69	10.69	1.67
Caribbean	38.79	27.27	36.49	13.32	1.54
Central American	41.96	30.77	35.06	12.30	1.48
Chilean	39.53	27.27	35.85	12.85	1.52
Spain Central	39.46	27.27	36.19	13.10	1.53
Spain Northern	37.92	25.00	36.20	13.10	1.53
Spain Southern	42.30	30.77	36.48	13.31	1.54
Mexican	38.44	26.14	36.17	13.09	1.53
Rioplátense	37.18	25.00	34.62	11.98	1.46

Table 8. WER (%): Alexa, US, Female.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	44.06	28.57	39.53	15.62	1.73
Canary	51.19	39.23	42.14	17.76	3.49
Caribbean	49.33	37.50	40.39	16.31	2.02
Central American	46.79	33.33	40.55	16.44	2.05
Chilean	49.46	40.00	40.01	16.00	2.20
Spain Central	40.33	23.08	40.01	16.00	1.77
Spain Northern	39.17	23.08	39.47	15.58	1.67
Spain Southern	48.85	37.50	39.68	15.75	2.04
Mexican	43.00	27.27	40.09	16.07	1.70
Rioplátense	44.55	30.77	39.66	15.73	1.68

Table 9. WER (%): Alexa, US, Male.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	47.91	35.71	40.82	16.66	0.75
Canary	48.85	36.36	39.49	15.59	2.01
Caribbean	47.01	33.33	41.58	17.29	1.05
Central American	48.37	36.36	40.06	16.05	1.26
Chilean	46.73	33.33	40.76	16.62	1.27
Spain Central	47.26	33.33	41.31	17.06	0.97
Spain Northern	46.01	30.77	41.22	16.99	0.46
Spain Southern	51.73	40.00	41.34	17.09	0.46
Mexican	48.93	33.33	42.05	17.68	0.38
Rioplátense	43.58	28.57	39.96	15.97	0.86

Table 10. WER (%): Whisper (Base), Female.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	25.72	15.38	38.38	14.73	1.68
Canary	18.08	10.00	25.23	6.37	2.09
Caribbean	26.05	16.67	32.66	10.66	1.63
Central American	22.34	14.29	27.39	7.50	1.39
Chilean	24.71	16.67	35.61	12.68	1.96
Spain Central	20.85	12.50	39.46	15.57	1.75
Spain Northern	21.69	14.29	32.98	10.87	1.08
Spain Southern	25.52	16.67	33.79	11.42	1.74
Mexican	25.67	18.18	30.81	9.49	0.32
Rioplátense	21.18	14.29	38.43	14.77	1.40

Table 11. WER (%): Whisper (Base), Male.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	22.71	14.29	53.11	28.21	0.97
Canary	21.78	16.67	23.24	5.40	1.18
Caribbean	26.46	16.67	37.24	13.87	0.94
Central American	20.70	14.29	28.99	8.40	0.91
Chilean	21.06	14.29	27.09	7.34	0.85
Spain Central	21.02	14.29	27.77	7.71	0.65
Spain Northern	17.95	10.00	37.28	13.90	0.41
Spain Southern	64.10	57.14	63.37	40.16	0.71
Mexican	26.35	16.67	45.14	20.37	0.41
Rioplátense	20.75	14.29	26.95	7.26	0.58

Table 12. WER (%): Whisper (Large-v2), Female.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	10.71	0.00	42.35	17.94	1.85
Canary	6.98	0.00	16.46	2.71	1.36
Caribbean	12.52	0.00	130.09	169.22	6.51
Central American	11.71	0.00	134.44	180.75	6.81
Chilean	8.23	0.00	19.82	3.93	1.09
Spain Central	15.08	0.00	222.53	495.22	9.86
Spain Northern	8.61	0.00	109.80	120.57	3.58
Spain Southern	8.67	0.00	22.04	4.86	1.13
Mexican	8.38	0.00	36.90	13.62	0.67
Rioplátense	8.37	0.00	54.07	29.24	1.96

Table 13. WER (%): Whisper (Large-v2), Male.

Accent	Mean	Median	Stdev	Variance	CI 95%
Andean	7.15	0.00	19.35	3.74	0.35
Canary	5.87	0.00	11.79	1.39	0.60
Caribbean	9.44	0.00	57.86	33.47	1.46
Central American	7.01	0.00	35.94	12.92	1.13
Chilean	7.17	0.00	18.06	3.26	0.56
Spain Central	6.60	0.00	30.57	9.34	0.72
Spain Northern	7.57	0.00	42.29	17.89	0.77
Spain Southern	21.44	8.33	34.54	11.93	0.63
Mexican	9.18	0.00	24.12	5.82	0.44
Rioplátense	6.45	0.00	19.59	3.84	0.42

Table 14 displays the gender differences for each variant of Alexa and Whisper analyzed by weighted means of WER. In our study, we employ a weighted arithmetic mean to analyze gender differences for each variant of Alexa and Whisper. This weighting approach assigns weights to each value based on the number of elements available for each option, thus ensuring a fair and representative assessment of the observed differences.

Table 14. Weighted mean WER (%) by gender.

Gender	Alexa MX	Alexa ES	Alexa US	Whisper Base	Whisper Large-v2
Female	39.30	36.92	44.72	24.76	9.42
Male	42.54	39.61	47.60	31.11	9.70

For better visualization of the results obtained, the following figures are presented, showing a comparison of the WER means and WER medians for females and males for each of the analyzed accents. Figure 1 illustrates the comparison of means for females. Figure 2 shows the comparison of means for males. Figures 3 and 4 display comparisons of medians for females and males, respectively. Additionally, in Figure 5, the weighted mean by gender is shown for each of the analyzed variants of Alexa and Whisper.

All data and results obtained during the research are available in the project repository. These datasets include transcribed and correct texts and WER for each phrase. Other error measures such as Character Error Rate (CER), Match Error Rate (MER), Word Information Lost (WIL), and Word Information Preserved (WIP) are also included in the datasets. The data are divided based on the tool used, i.e., Alexa, Whisper Base, or Whisper Large-v2, and within these categories, the data is further divided by gender and accent [41].

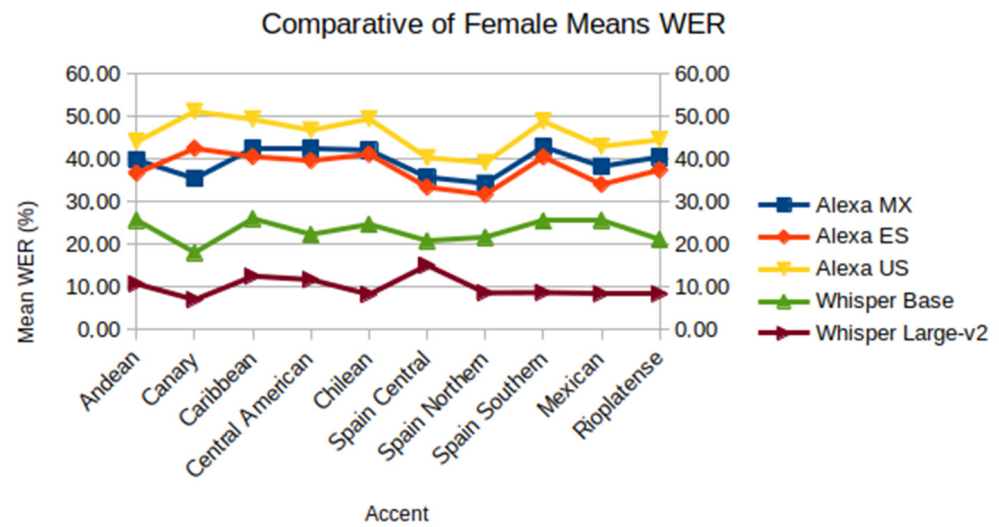


Figure 1. Comparative of females means WER.

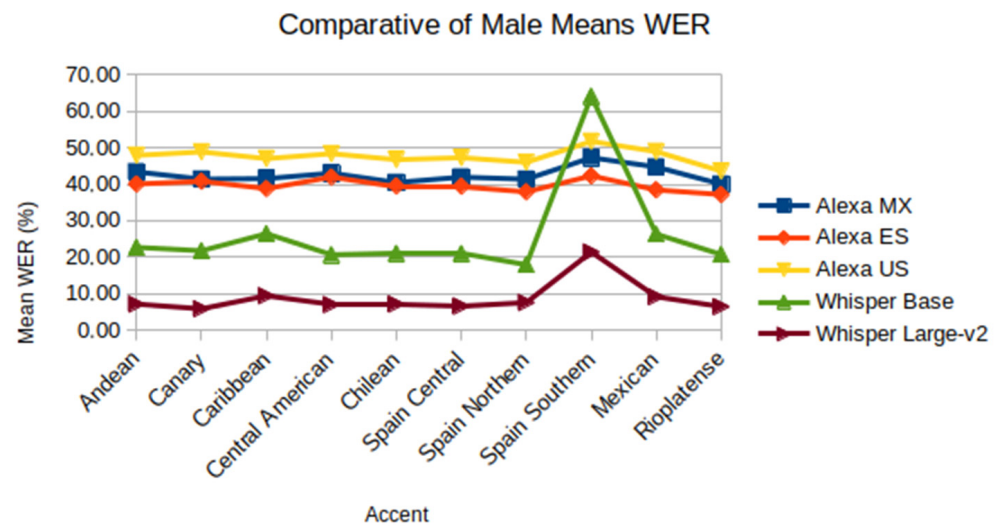


Figure 2. Comparative of males means WER.

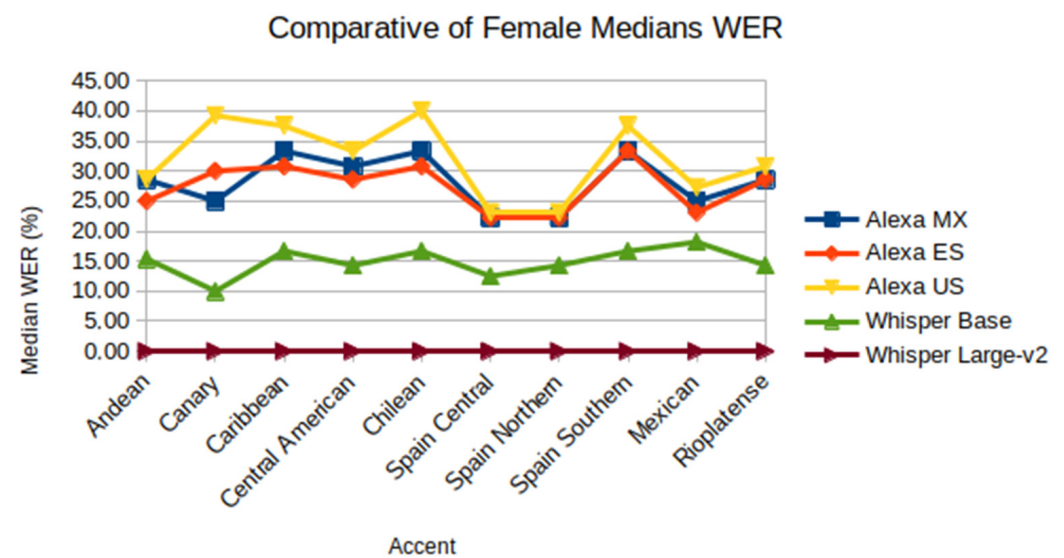


Figure 3. Comparative of females' medians WER.

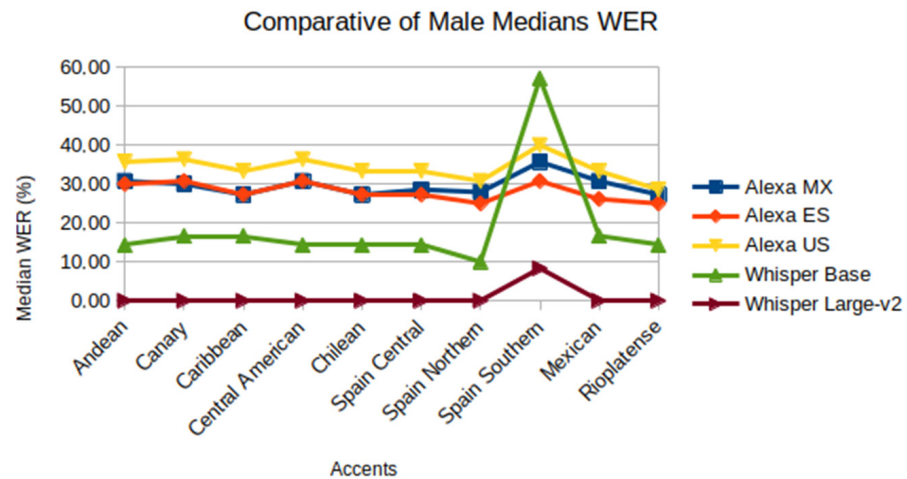


Figure 4. Comparative of males’ medians WER.

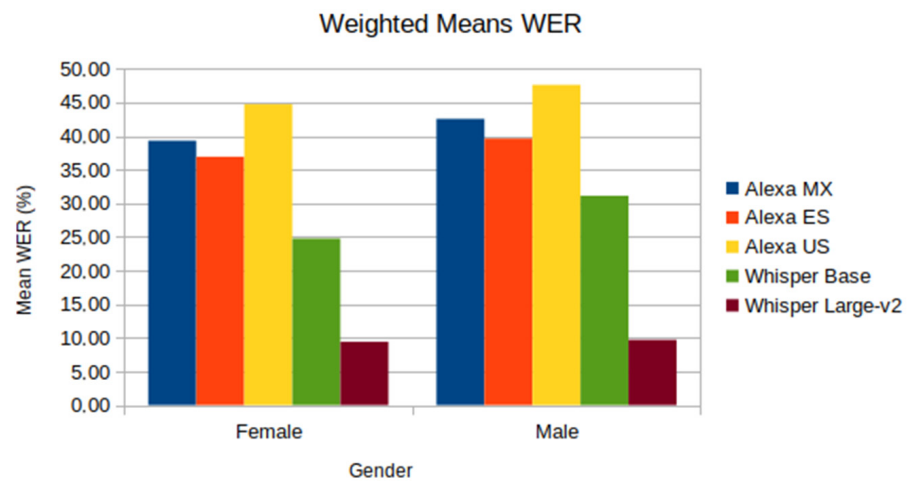


Figure 5. Comparative of weighted means.

5. Discussion

After analyzing the results, it is observed that, as a general rule, the Alexa variant for U.S. Spanish, here identified as Alexa US, performs the worst among the three Alexa variants for any of the analyzed accents. In the case of the other two Alexa variants, for Mexican Spanish and Spanish from Spain, identified here as Alexa MX and Alexa ES, respectively, the comparison of means reveals a similar performance, with a slight improvement in Alexa ES compared to Alexa MX, except for the case of females from the Canary Islands, where Alexa MX performs better. In the case of median comparisons, we see virtually the same result, except for several cases with the same median for both Alexa MX and Alexa ES. Additionally, there are two cases where Alexa MX outperforms Alexa ES, and these cases are for both genders of the Canary accent, although the difference is more pronounced for females.

When comparing the results obtained in the different Alexa variants with the results of Whisper (Base), it is generally observed that Whisper performs significantly better than Alexa for any of the analyzed accents, except for the variant spoken in the Southern Iberian Peninsula and for males. Median values for Alexa are around 30%, while Whisper has about 15% for WER. Generally, Whisper (Base) makes approximately half the WER errors as Alexa with the Spanish language.

In the case of Whisper (Large-v2), it is observed that the results significantly improve compared to Whisper (Base), with a mean below 10% for both women and men. Particularly indicative are the median values of Whisper (Large-v2) for both genders, as they achieve medians of 0% for all cases except for the Spanish of the Southern Iberian Peninsula.

These data reveal a clearly identifiable outlier for Whisper, both for the Base model and the Large-v2 model, in the accent of Spanish from the southern part of the Iberian Peninsula when spoken by males. After analyzing the Common Voice 14 dataset, it is observed that thousands of contributions could be from the same person, which may have influenced the final result of the analysis for this accent.

Taking into account that the selected sample for the male gender and the accent of the Southern Iberian Peninsula was 30,698 audio segments for Whisper (Base), 11,600 for Whisper (Large-v2), and 2150 for Alexa, it is possible that this influence did not affect the analysis of Alexa as much as it did for Whisper. If the analysis results for females had also deteriorated for this accent, it would not have been assumed that something was affecting the male results.

In the case of using the same random sample for Whisper (Base) for the Southern Iberian Peninsula accent as used for Alexa, the results obtained are shown in Table 15. Clearly, there is bias in the dataset affecting the results. The median values for Whisper (Base) are 16.67%, which is closer to the results obtained for females for this accent. For Whisper (Large-v2), Similar results were obtained, with a median of 0.00%, which is quite similar to the results obtained for females. These data can be seen in Table 16.

Table 15. Whisper (Base): South of Spain.

Gender	Mean	Median	Stdev	Variance	CI 95%
Female	25.52	16.67	33.79	11.42	1.74
Male (2150)	26.92	16.67	49.60	24.60	2.10

Table 16. Whisper (LargeV2): South of Spain.

Gender	Mean	Median	Stdev	Variance	CI 95%
Female	8.67	0.00	22.04	4.86	1.13
Male (2150)	9.10	0.00	19.62	3.85	0.83

If the weighted means obtained previously and shown in Table 14 and Figure 5 are recalculated using this correction for this accent, the data obtained is shown in Table 17 and Figure 6.

Table 17. Weighted means WER by gender.

Gender	Alexa MX	Alexa ES	Alexa US	Whisper Base	Whisper Large-v2
Female	39.30	36.92	44.72	24.76	9.42
Male	42.54	39.61	47.60	23.57	7.84

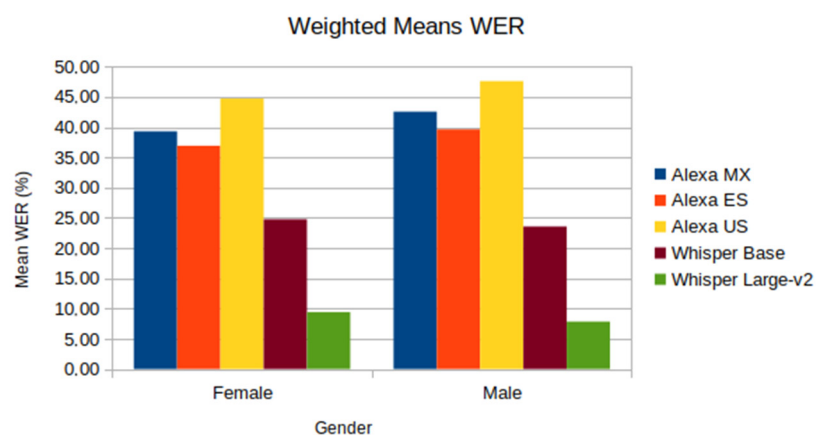


Figure 6. Comparative of weighted means WER by gender.

Let us analyze the weighted means shown in Table 17 and Figure 6. We can observe that in all cases, for the same type of Alexa variant, female voices are slightly better recognized than male voices, meaning the mean WER is lower for women than for men. On the other hand, for Whisper, the opposite is true—the mean WER is slightly lower for men than for women.

More specifically, we see that in Alexa, the difference between the mean WER for female and male voices is 3.24 pp (percentage points), 2.69 pp, and 2.88 pp, respectively, for Alexa MX, Alexa ES, and Alexa US, in favor of female voices. In the case of Whisper, these differences are 1.19 pp and 1.58 pp for the Base and Large-v2 models, respectively, but in favor of male voices. Table 18 and Figure 7 show the average data for each of the Alexa and Whisper variants, regardless of gender.

Table 18. Weighted means WER by accent.

Gender	Alexa MX	Alexa ES	Alexa US	Whisper Base	Whisper Large-v2
Andean	41.66	38.44	46.05	23.16	7.68
Canary	39.76	41.29	49.49	20.76	6.18
Caribbean	41.91	39.54	47.98	26.38	10.06
Central American	42.79	41.01	47.72	21.16	8.31
Chilean	41.08	40.13	47.74	21.95	7.43
Spain Central	38.99	36.58	43.96	20.98	8.47
Spain Northern	37.85	34.81	42.59	18.34	7.82
Spain Southern	45.49	41.57	50.57	26.86	9.05
Mexican	41.50	36.26	45.97	26.06	8.78
Rioplatense	40.26	37.32	44.07	20.86	6.95

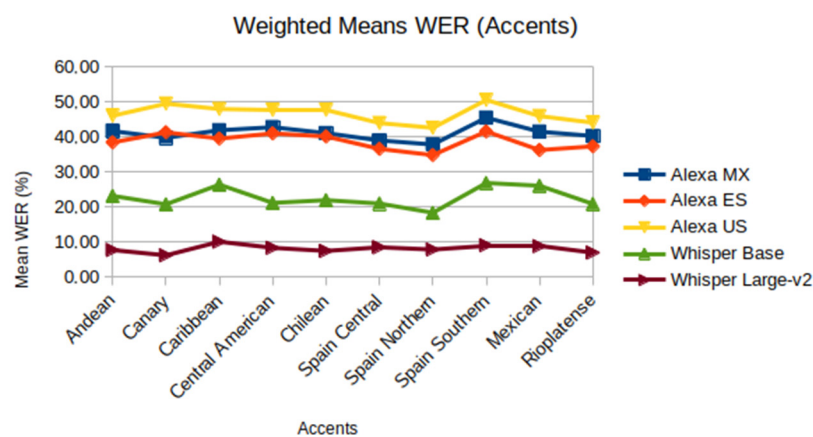


Figure 7. Comparative of weighted means WER by Accent.

In these weighted mean data for each accent, each variant of Alexa and Whisper was analyzed, and it can be observed that the Alexa US model performs the worst in all cases compared to the other variants. In all other cases, the Alexa ES variant is slightly better than the Alexa MX variant, except for the Canarian accent, where it is the opposite.

It is noticeable that in some instances, the difference in the weighted means WER by accent; for example, for Alexa, it can reach up to 6.76 pp. This is the case when comparing the Spanish Northern and Spanish Southern accents. For Alexa MX, the difference between the Northern and Southern Spanish accents can be as high as 7.64 pp. For Alexa US, 7.98 pp between the same accents. These differences are observed across the various accents as well.

In Whisper (Base), approximately 8 pp differences are observed between the Northern Spanish accent and the Caribbean, Mexican, and Southern Spanish accents. In Whisper (Large-v2), a difference of around 4 pp between the Canarian and Caribbean accents is noted.

When analyzing the confidence intervals (CI 95%) calculated for each of the available options, it is observed that, overall, the sample fits appropriately and is accurately repre-

sented by these statistical measures. A wider confidence interval is notable in the case of Whisper Large-v2 for females, especially for Caribbean, Central American, and Central Spanish accents, although it is also significantly high for the Northern Spanish accent, albeit to a lesser extent. Although these values encompass a broader probable range than desired, they are not sufficiently large to invalidate this portion of the results. In the remaining cases, for both genders and all analyzed tools, the confidence intervals are below 2, and even in many cases below 1, with a slight exception for Canary Islands females and Alexa variants, where we obtain 2.88 for MX, 3.13 for ES, and 3.49 for US. However, considering that the mean WER values for these variants are 35.46, 42.52, and 51.19, respectively, it is concluded that these values are still significant.

In previous studies, the analysis of gender and accent bias within the same language has been explored. For instance, examining the DeepSpeech (STT) model using data from the Mozilla Common Voice project revealed bias between different English variants, specifically US and Indian English. However, the study concluded that there is no significant evidence of gender bias in the DeepSpeech model [33].

Another study that assessed YouTube's ASR system in transcribing voice-to-text in platform-uploaded videos estimated gender bias as disadvantageous to women compared to men. It also identified accent bias among different studied variants, particularly disadvantaging the Scottish accent [31].

A separate study demonstrated that underrepresenting, for example, the female gender in the dataset used to train an ASR system results in a higher WER for that gender, indicating the presence of gender bias [34].

6. Conclusions

Based on the observed data, it can be asserted that Alexa performs better in recognizing female speech in Spanish, while, conversely, Whisper exhibits better performance for male speech. However, with a mean difference of 2.94 pp for Alexa (favoring females) and 1.39 pp for Whisper (favoring males), we consider these differences not significant enough to conclude the existence of gender bias in Alexa and Whisper for the Spanish language or at least not a bias that significantly influences the everyday functionality of these tools. This small difference could be attributed to sample error, especially considering the WER for both Alexa and Whisper.

More significant differences are observed regarding accents, reaching up to 8 pp in some cases. As a general rule for Alexa, there appears to be a bias in favor of the Northern Spanish accent and against the Southern Spanish accent primarily, as well as other accents such as Caribbean, Central American, and Canarian. For Whisper (Base), there seems to be a potential bias in favor of the northern accent compared to the southern accent in Spain. In the case of Whisper (Largev2), this bias is mainly in favor of the Canarian accent as opposed to the Caribbean accent.

An interesting fact regarding the data obtained with Whisper (Large-v2) is that the Canarian is precisely the most accurately recognized despite being the dialectal variant with the fewest speakers. The Canary Islands have just under 2.2 million inhabitants [42], compared to the nearly 500 million people with Spanish as their native language [43].

Concerning the three Alexa models tailored for Mexican Spanish, Spanish from Spain, and Spanish from the United States, there appears to be no compelling reason to maintain these three models as separate entities for the Spanish language. It would be preferable to the optimized model, similar to the approach taken by Whisper. Subsequently, if the goal is to enable Alexa to speak in different dialects, employing a TTS system trained specifically for each accent in question would be more suitable.

One of the fundamental limitations of this project arises from the fact that the audio files included in the Common Voice dataset are derived from segments of read texts, in contrast to the natural way of interacting with a voice-activated virtual assistant, which typically involves conversations. Another limitation is that the number of available audio

fragments is not comparable between those read by males and females, and a similar issue exists with the distribution across different accents.

Concerning future research directions, an obvious avenue would be to extend the study to other languages and their respective dialects, enabling a more comprehensive comparison. Another potential research extension could involve a specific comparison for Alexa, examining the speech-recognition error rate when using a set of predefined words versus a set of non-defined words, as addressed in this study. In many cases, Alexa Skills employs predefined sets of words to facilitate the interaction flow.

Analyzing automatic speech recognition with other datasets and for the various languages available for each system would enable us to determine whether progress is genuinely being made in eliminating gender bias in ASR systems or if, in this case, these results are specific to the Spanish language and these two particular tools.

Author Contributions: Conceptualization, methodology, formal analysis, data curation, E.N.-G.; writing—original draft preparation, E.N.-G.; writing—review and editing, H.S.D.-K.-N. and C.S.G.-G.; supervision C.S.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was co-financed by the Canary Islands Agency for Research, Innovation, and Information Society of the Ministry of Economy, Knowledge and Employment and by the European Social Fund (ESF) Integrated Operational Program of the Canary Islands 2014–2020, Axis 3 Priority Topic 74 (85%). This work has been supported partially by the PERGAMEX ACTIVE project, Ref. RTI2018-096986-B-C32, funded by the Ministry of Science and Innovation. Spain. This work has been supported partially by the PLEISAR-Social project, Ref. PID2022-136779OB-C33, funded by the Ministry of Science and Innovation. Spain. This work has been partially supported by the COEDUIN project, Ref. 2020EDU08, funded by Fundación Caja Canarias and Fundación La Caixa.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10152506>, reference number 10.5281/zenodo.10152506.

Acknowledgments: We want to thank the Canary Islands Agency for Research, Innovation, and Information Society of the Ministry of Economy, Knowledge and Employment and the European Social Fund (ESF) Integrated Operational Program of the Canary Islands 2014–2020, Axis 3 Priority Topic 74 (85%), the PERGAMEX ACTIVE project, Ref. RTI2018-096986-B-C32, funded by the Ministry of Science and Innovation. Spain and the PLEISAR-Social project, Ref. PID2022-136779OB-C33, funded by the Ministry of Science and Innovation. Spain, the COEDUIN project, Ref. 2020EDU08, funded by Fundación Caja Canarias and Fundación La Caixa, for their support.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Beirl, D.; Yuill, N.; Rogers, Y. Using Voice Assistant Skills in Family Life. June 2019. Available online: <https://repository.isls.org/handle/1/1750> (accessed on 16 November 2023).
2. Porcheron, M.; Fischer, J.E.; Reeves, S.; Sharples, S. Voice Interfaces in Everyday Life. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Montreal, QC, Canada, 21–26 April 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1–12. [CrossRef]
3. Këpuska, V.; Bohouta, G. Next-Generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In Proceedings of the 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 8–10 January 2018; pp. 99–103. [CrossRef]
4. Ford, M.; Palmer, W. Alexa, Are You Listening to Me? An Analysis of Alexa Voice Service Network Traffic. *Pers. Ubiquitous Comput.* **2019**, *23*, 67–79. [CrossRef]
5. Home Assistant. Available online: <https://www.home-assistant.io/> (accessed on 16 November 2023).
6. Vázquez, J.J.P.; Ortiz, C.A.N.; Cordero, M.P.O.; León, P.A.P.; Orellana, P.C. Evaluación del reconocimiento de voz entre los servicios de Google y Amazon aplicado al Sistema Integrado de Seguridad ECU 911. *Revista Tecnológica ESPOL* **2021**, *33*, 2. [CrossRef]
7. Whisper. Python. 2022. Reprint, OpenAI. Available online: <https://github.com/openai/whisper> (accessed on 16 November 2023).

8. Coqui-Ai/STT: STT—The Deep Learning Toolkit for Speech-to-Text. Training and Deploying STT Models Has Never Been So Easy. Available online: <https://github.com/coqui-ai/STT> (accessed on 16 November 2023).
9. Seaborn, K.; Urakami, J. Measuring Voice UX Quantitatively: A Rapid Review. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21), Yokohama, Japan, 8–13 May 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 1–8. [CrossRef]
10. Worldometer. Worldometer—Real Time World Statistics. Available online: <http://www.worldometers.info/> (accessed on 16 November 2023).
11. Ethnologue (Free Dev). What Are the Top 200 Most Spoken Languages? Available online: <https://www.ethnologue.com/insights/ethnologue200/> (accessed on 16 November 2023).
12. Aguirre, A.; Manasía, N. Derechos humanos de cuarta generación: Inclusión social y democratización del conocimiento. *Télématique* **2015**, *14*, 2–16.
13. UNESCO. Education for All 2000–2015: Achievements and Challenges | Global Education Monitoring Report. Available online: <https://www.unesco.org/gem-report/en/efa-achievements-challenges> (accessed on 19 April 2022).
14. Costa-jussà, M.R.; Basta, C.; Gállego, G.I. Evaluating Gender Bias in Speech Translation. *arXiv* **2022**, arXiv:2010.14465.
15. Reid, K.; Williams, E.T. Common Voice and accent choice: Data contributors self-describe their spoken accents in diverse ways. In Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '23), Boston, MA, USA, 30 October–1 November 2023; Association for Computing Machinery: New York, NY, USA, 2023; pp. 1–10. [CrossRef]
16. Nguējajio, M.K.; Washington, G. Hey ASR System! Why Aren't You More Inclusive? In *HCI International 2022—Late Breaking Papers: Interacting with eXtended Reality and Artificial Intelligence*; Chen, J.Y.C., Fragomeni, G., Degen, H., Ntoa, S., Eds.; Lecture Notes in Computer Science; Springer Nature: Cham, Switzerland, 2022; pp. 421–440. [CrossRef]
17. Feng, S.; Kudina, O.; Halpern, B.M.; Scharenborg, O. Quantifying Bias in Automatic Speech Recognition. *arXiv* **2021**, arXiv:2103.15122.
18. Markl, N. Language Variation, Automatic Speech Recognition and Algorithmic Bias. Ph.D. Thesis, The University of Edinburgh, Edinburgh, UK, 2023. [CrossRef]
19. Wassink, A.B.; Gansen, C.; Bartholomew, I. Uneven success: Automatic speech recognition and ethnicity-related dialects. *Speech Commun.* **2022**, *140*, 50–70. [CrossRef]
20. Markl, N. Language variation and algorithmic bias: Understanding algorithmic bias in British English automatic speech recognition. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), Seoul, Republic of Korea, 21–24 June 2022; Association for Computing Machinery: New York, NY, USA, 2022; pp. 521–534. [CrossRef]
21. Vorvoreanu, M.; Zhang, L.; Huang, Y.-H.; Hilderbrand, C.; Steine-Hanson, Z.; Burnett, M. From Gender Biases to Gender-Inclusive Design: An Empirical Investigation. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19), Glasgow, UK, 4–9 May 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–14. [CrossRef]
22. Breslin, S.; Wadhwa, B. Gender and Human-Computer Interaction. In *The Wiley Handbook of Human Computer Interaction*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2018; pp. 71–87. [CrossRef]
23. Mulas, C.M. Speech Signals Feature Extraction Model for a Speaker's Gender and Age Identification System. Ph.D. Thesis, E.T.S. de Ingenieros Informáticos (UPM), Madrid, Spain, 2014. Available online: <https://oa.upm.es/33121/> (accessed on 28 March 2024).
24. Latinus, M.; Taylor, M.J. Discriminating Male and Female Voices: Differentiating Pitch and Gender. *Brain Topogr* **2012**, *25*, 194–204. [CrossRef] [PubMed]
25. Titze, I.R. Physiologic and acoustic differences between male and female voices. *J. Acoust. Soc. Am.* **1989**, *85*, 1699–1707. [CrossRef] [PubMed]
26. Chela-Flores, G. La División Dialectal Del Español. In *Dialectología Hispánica/The Routledge Handbook of Spanish Dialectology*; Routledge: Oxford, UK, 2022.
27. Mozilla. Mozilla Common Voice. Available online: <https://commonvoice.mozilla.org/> (accessed on 16 November 2023).
28. Bias in AI: What It Is, Types, Examples & 6 Ways to Fix It in 2023. Available online: <https://research.aimultiple.com/ai-bias/> (accessed on 14 November 2023).
29. De Oliveira, C.B.; Amaral, M.A. A discourse analysis of interactions with Alexa virtual assistant showing reproductions of gender bias. *Clepsydra. Rev. Int. De Estud. De Género Y Teoría Fem.* **2022**, *23*, 37–58. [CrossRef]
30. Oliveira, C.B.; Amaral, M.A. An Analysis of the Reproduction of Gender Bias in the Speech of Alexa Virtual Assistant. In Proceedings of the XIII Congress of Latin American Women in Computing, San José, Costa Rica, 25–29 October 2021.
31. Tatman, R. Gender and Dialect Bias in YouTube's Automatic Captions. In Proceedings of the First ACL Workshop on Ethics in Natural Language Processing, Valencia, Spain, 4 April 2017; Association for Computational Linguistics: Valencia, Spain, 2017; pp. 53–59. [CrossRef]
32. Sun, T.; Gaut, A.; Tang, S.; Huang, Y.; ElSherief, M.; Zhao, J.; Mirza, D.; Belding, E.; Chang, K.-W.; Wang, W.Y. Mitigating Gender Bias in Natural Language Processing: Literature Review. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019; Korhonen, A., Traum, D., Màrquez, L., Eds.; Association for Computational Linguistics: Florence, Italy, 2019; pp. 1630–1640. [CrossRef]

33. Meyer, J.; Rauchenstein, L.; Eisenberg, J.D.; Howell, N. Artie Bias Corpus: An Open Dataset for Detecting Demographic Bias in Speech Applications. In Proceedings of the Twelfth Language Resources and Evaluation Conference, Marseille, France, 11–16 May 2020; European Language Resources Association: Marseille, France, 2020; pp. 6462–6468. Available online: <https://aclanthology.org/2020.lrec-1.796> (accessed on 3 July 2023).
34. Garnerin, M.; Rossato, S.; Besacier, L. *Investigating the Impact of Gender Representation in ASR Training Data: A Case Study on Librispeech*; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 86–92. [CrossRef]
35. Amazon Alexa. Amazon Alexa Voice AI | Alexa Developer Official Site. Available online: <https://developer.amazon.com/enUS/alexa.html> (accessed on 16 November 2023).
36. Introducing Whisper. Available online: <https://openai.com/research/whisper> (accessed on 16 November 2023).
37. Selenium. Available online: <https://www.selenium.dev/> (accessed on 16 November 2023).
38. Eren, G.; The Coqui TTS Team. Coqui TTS. Python. January 2021. [CrossRef]
39. Amazon Alexa. Slot Type Reference | Alexa Skills Kit. Available online: <https://developer.amazon.com/en-US/docs/alexa/customskills/slot-type-reference.html> (accessed on 16 November 2023).
40. Morris, A.C.; Maier, V.; Green, P. From WER and RIL to MER and WIL: Improved Evaluation Measures for Connected Speech Recognition. In Proceedings of the Interspeech, Jeju Island, Republic of Korea, 4–8 October 2004; pp. 2765–2768. [CrossRef]
41. Nacimiento-Garcia, E. Menceybencomo/Comparative-Analysis-of-Gender-and-Accent-Biases-in-Alexa-and-Whisper-for-the-Spanish-Language: Paper Version. [CrossRef]
42. Instituto Canario de Estadística. Demografía. Available online: <https://www.gobiernodecanarias.org/istac/estadisticas/demografia/index.html> (accessed on 16 November 2023).
43. Centro Virtual Cervantes. CVC. Anuario 2023. Informe 2023. El Español en Cifras. Instituto Cervantes. Available online: https://cvc.cervantes.es/lengua/anuario/anuario_23/informes_ic/p01.htm (accessed on 16 November 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.