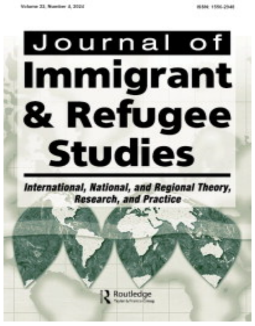

Journal of Immigrant & Refugee Studies

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/wimm20

Machine Discriminating: Automated Speech Recognition Biases in Refugee Interviews

Tuba Bircan & Duha Ceylan

To cite this article: Tuba Bircan & Duha Ceylan (21 May 2024): Machine Discriminating: Automated Speech Recognition Biases in Refugee Interviews, Journal of Immigrant & Refugee Studies, DOI: 10.1080/15562948.2024.2355613

To link to this article: https://doi.org/10.1080/15562948.2024.2355613

JOURNAL OF IMMIGRANT & REFUGEE STUDIES
https://doi.org/10.1080/15562948.2024.2355613

Machine Discriminating: Automated Speech Recognition Biases in Refugee Interviews

Tuba Bircan and Duha Ceylan

BRISPO, Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium

ABSTRACT
This study scrutinizes Automated Speech Recognition (ASR) software, a powerful instrument to expedite the translation process, and their unintended bias against Arabic speakers, particularly refugees. We propose that pre-existing biases in ASR training data reflect societal prejudices, leading to orientalist and Islamophobic misrepresentations. We used four ASR tools to transcribe interviews with Arabic-speaking refugee women, employing ideological textual analysis to detect biases. Our findings indicate that ASR algorithms may inadvertently associate Arabic speakers with conflict, war, religion, and narrow down Arab identities to Islam-centric representations. Acknowledging these biases is essential for fostering a more equitable and culturally sensitive technological environment.

KEYWORDS
Automated transcription; ideological textual analysis; bias in machine learning; refugee interviews; automated speech recognition

Introduction

In the field of qualitative research, a domain deeply rooted in dialogical and interview-based methodologies, the transcription of verbal discourse into textual data emerges as a pivotal yet oft-overlooked component in scholarly discourse (Davidson, 2009). Historically, qualitative research methods necessitated the concurrent documentation of interviews, a practice laden with the potential for errors and inefficiency. The advent of voice recorders signified a paradigm shift in data collection, allowing researchers to capture the full breath of verbal exchanges, thereby minimizing the risk of data loss or misinterpretation.

Despite this technological progression, the transcription of recorded dialogues remains a laborious endeavor, frequently demanding thrice the duration of the original recordings for accurate transcription. The recent emergence of automated transcription software has heralded a more efficient approach to this protracted task, finding diverse applications in language learning, accessibility enhancement, psychosocial support, and streamlined documentation processes (Choe et al., 2019). Moreover, advancements in machine learning (ML) have catalyzed the development of innovative methods for sociological knowledge-driven textual analysis (Németh et al., 2020), thereby enriching the interpretative depth of qualitative data.

Nevertheless, the burgeoning utilization of Automated Speech Recognition (ASR) systems within research contexts is not without its caveats. A primary concern centers around the inherent biases of ASR systems, which are inextricably linked to the training data used in their algorithmic developments. Studies have highlighted a predilection of ASR systems for certain accents or dialects, leading to a disproportionate error rate amongst speakers of varied linguistic backgrounds (Koenecke et al., 2020). Such biases can engender misrepresentations and misunderstandings, particularly in sensitive settings, such as refugee interviews.

CONTACT Tuba Bircan tuba.bircan@vub.be BRISPO, Department of Sociology, Vrije Universiteit Brussel, Brussels, Belgium

In the sphere of migration governance, ASR systems hold promise for digitizing various processes, including visa applications, interviews, and language translation. However, the implementation of these systems necessitates careful consideration due to the possibility of biases and inaccuracies that could negatively affect vulnerable individuals. The role of algorithmic decision-making in these processes, critiqued by Eubanks (2018) and Molnar (2023), reflects a broader concern about systemic biases and the risk of oversimplified technological enforcement in migration management. This parallels the specific biases in ASR systems identified in interviewing Arabic-speaking refugees, underscoring the need for ethical technology application in migration. The importance of addressing ASR and AI biases for ongoing research to understand and mitigate ASR system limitations, particularly in migration contexts (Katzenbach & Ulbricht, 2019; Koenecke et al., 2020; Latonero & Kift, 2018).

This study aims to scrutinize the ideological biases and stereotypes that may arise from the employment of ASR transcription software in interviews with refugee populations, particularly focusing on Arabic-speaking refugee cohort. Through ideological textual analysis (ITA), the research endeavors to unravel biases pertaining to religious, societal, and linguistic dimensions. As in-depth comprehension of these biases is vital for the precise interpretation of ASR outputs and to mitigate any potential harm to vulnerable groups. As ASR systems become increasingly integral to research methodologies, an acute awareness of their inherent biases is essential for their ethical and responsible utilization.

The ensuing sections of this article will commence with an exploration of the conceptual facets surrounding studies of Arabic-speaking migrant communities. This will be followed by a review of the extant literature on the challenges associated with the application of ASR technology, a detailed exposition of our data collection and methodological approach, an analysis of our findings within the context of conceptual frameworks, and finally, a discussion encompassing our findings, the limitations of our study, and potential avenues for future research.

Studying Arabic-speaking migrant communities

The increase in refugee movements, especially from Syria to Europe and nearby regions, has intensified research focusing on refugees and migrants, highlighting the importance of qualitative methodologies using Arabic-speaking interviewers (Al-Amer et al., 2016). Transcribing interviews in Arabic remains a time-consuming task, even for Arabic-proficient researchers, significantly slowing down research processes (Aloudah, 2022). The use of computer-aided transcription tools offers a solution to these challenges but necessitates careful consideration of their suitability, especially for interviews with predominantly Syrian, Arabic-speaking, and Muslim participants, who often have complex migrant or refugee experiences.

In this milieu, the imperative for decolonizing methodological becomes pronounced, especially in research involving vulnerable populations (Thambinathan & Kinsella, 2021). The pervasive influence of Eurocentric paradigms extends beyond mere methodological approaches (Held, 2019), permeating the very instruments utilized for data collection and subsequent data analysis. A critical understanding of orientalism and Islamophobia, as manifest in public and political discourse, is imperative to contextualize the positioning of these communities (Ali, 2017; Feldman & Medovoi, 2016). Orientalism, a concept first explicated by Said (1979) delineates the West's intellectual and cultural patronage over the East, fostering perceptions steeped in paternalism and ethnocentrism. This discourse, in turn, fuels the proliferation of Islamophobia, here conceptualized as the amalgamation of negative stereotypes and biases against Islam and its adherents, rooted in orientalist caricatures.

Therefore, examining ideological biases in automated transcription tools extends beyond technical analysis to a critical exploration of potential orientalist and Islamophobic biases. This perspective is essential for understanding how such biases affect the portrayal and interpretation of Arabic-speaking communities. Our study integrates this critical approach with a broader examination of the socio-cultural and ideological contexts surrounding migrant communities, aiming to uncover how these influences shape transcription outcomes.

Understanding the ASR

ASR represents a technological foray into the sphere of computational linguistics, designed to identify and interpret spoken language captured *via* microphones or telephonic devices. At its core, ASR employs computational algorithms to discern spoken language, effectively translating the acoustic signals into coherent strings of words. ASR systems are adept at transforming the acoustic realization of the speech into a sequenced array of linguistic units or words. This transformation can be achieved through either conventional ASR methodology or *via* advanced deep learning techniques.

In the traditional paradigm, ASR algorithms generate features from the audio inputs and subsequently engage an acoustic model (AM) to match phonemes. This is followed by a language model (LM), which utilizes probability distributions to predict words and their sequential arrangement. Conversely, deep learning approaches integrate the functions of AM and LM. The alignment of training data with test conditions is an important factor influencing ASR performance. It is imperative to note that, akin to many other Artificial Intelligence (AI) processes, deep learning relies on the diversity and robustness of training corpora for any given language. These AI algorithms do not differentiate between factual data and user-generated content, thereby making the patterns within the training information a potential conduit for the transmission of initial biases.

The efficacy of an ASR system is conventionally evaluated using the Word Error Rate (WER), a metric that quantifies errors in an ASR generated transcript as a percentage of incorrect words, inclusive of insertions, deletions, and substitutions when contrasted with a standard transcription. It is noteworthy that WER primarily focuses on lexical accuracy, disregarding the semantic context of words. Studies have demonstrated that ASR systems exhibit superior performance in transcribing conversational telephone speech when assessed in terms of WER (Xiong et al., 2018). However, the semantic integrity of speech, as important as lexical accuracy, can lead to divergent connotations and interpretations, particularly when juxtaposed between spoken and written modalities (Collins et al., 2019).

Contemporary research has harnessed advanced computational methodologies, such as Natural Language Processing (NLP) for grammar error detection (Choe et al., 2019) and contextual language analysis within ASR outputs, addressing issues, such as racial bias (i.e., Koenecke et al., 2020; Tan et al., 2020). The challenges of addressing connotation problems, indicative of socio-political, cultural, or racial biases, differ significantly from mere error detection. The application of NLP in this context is notably more complex, focusing on pattern recognition and deep neural network-style ML methods, with existing research indicating inherent biases within NLP itself (Sun et al., 2019).

To elucidate, research has revealed gender biases in voice recognition technologies, with systems trained on male voices performing worse on female voices, emphasizing the importance of diverse training data for gender equitable recognition (May et al., 2019; Tatman & Kasten, 2017). Additionally, racial biases in ASR systems, demonstrated by lower accuracy for certain accents or dialects, point to ingrained biases (Koenecke et al., 2020). These biases compromise system utility and reinforce stereotypes, highlighting the need for careful, inclusive ASR development.

Accuracy, bias, and stereotyping

ASR algorithms have witnessed substantial advancements in voice-to-text accuracy, particularly for globally dominant languages like English (Xiong et al., 2018). Nevertheless, languages with lesser resources and smaller speaking populations, such as Arabic, encounter formidable challenges in attaining comparable accuracy levels (Eberhard et al., 2019). Arabic, with its complex morphology and extensive lexicon, presents considerable complexities in amassing adequate resources for effective acoustic, pronunciation, and language model training (Dhouib et al., 2022).

In the context of social science research, the deployment of automated transcription services raises critical issues concerning precision and algorithmic bias. Distinct from researcher bias, which pertains to methodological inclinations or interpretive biases, AI algorithms can manifest biases affecting voice-to-text accuracy and specificities of content (Blodgett et al., 2020; Garg et al., 2018). Such biases within AI algorithms predominantly originate from the preconceptions embedded in the training data and annotated speech corpora (Davidson et al., 2019; Kiritchenko & Mohammad, 2018). Investigations into gender and racial biases in AI systems have unveiled disparities and negative generalizations, highlighting the need for ethical considerations in AI development (Koenecke et al., 2020; May et al., 2019). For example, studies highlight gender bias in AI, where voice recognition systems often fare better with male than female voices, causing transcription errors for women (Tatman & Kasten, 2017). Racial bias is also evident, with lower recognition accuracy for African American Vernacular English speakers compared to Standard American English (Blodgett et al., 2020). These issues arise from the lack of diverse voices in training datasets, resulting in systems that underperform for certain demographic groups.

However, specific biases related to the Arabic language and its cultural nuances have received relatively scant attention. Stereotyping and prejudices, such as Islamophobia and orientalism, remain largely unexamined within the purview of ASR systems (Mozafari et al., 2020; Vidgen & Yasseri, 2020). While existing research on hate speech predominantly focuses mainly on user-generated texts, the biases inherent in speech-to-text algorithms are yet to be thoroughly explored. A critical assessment of contextual and conceptual biases in annotated speech is crucial, especially considering the implicit associations between Arabic input and Islam. Such associations between the Arabic language and Islam go beyond linguistic connections, reflecting deep cultural and religious significance. Arabic's link to Islamic texts often implies an automatic association with Islam, a blend of language with cultural and religious symbolism (Farghaly, 2010). This overlap may introduce biases in ASR systems, where Arabic expressions could evoke stereotypes or misconceptions based on Islamophobic attitudes (Allen, 2020; Awan, 2016).

This study endeavors to augment the current corpus of knowledge by employing ideological textual analysis to discern biases in automated transcription services, with a specific focus on Arabic input. It examines the performance of ASR algorithms with an emphasis on socio-cultural sensitivity. By elucidating biases inherent in ASR systems and deepening our understanding of socio-cultural dynamics, this research seeks to mitigate potential detrimental effects and foster equitable and accurate outcomes in sensitive contexts.

Theoretical framework

Our research integrates a deconstructive framework, drawing upon Foucault's analysis of societal power dynamics (1972), Van Dijk's insights on cognitive ideology (1998, 2013), and Said's post-colonial critique (1979), to scrutinize biases in ASR technology used with Arabic-speaking refugees. The deconstruction approach, initiated by Derrida (2001), is crucial for challenging binary oppositions and inherent hierarchies, enabling a detailed examination of ASR technologies' nuances. It reveals the subtle biases and power structures embedded within ASR outputs, highlighting how these may influence perceptions of Arabic-speaking refugees.

Foucault's discourse theory provides a lens to understand the power relations that shape knowledge production and technology development. Van Dijk's exploration of cognitive ideology complements this by analyzing how societal biases are internalized and replicated in technology, affecting the representation of minorities. Said's examination of orientalism (1979) adds a critical perspective on the cultural and historical biases that ASR technologies risk perpetuating, emphasizing the need for awareness in technological development and application.

Together, these theoretical perspectives offer a comprehensive framework for analyzing the complex interplay between technology, society, and entrenched biases. They guide our critique of ASR systems, advocating for the creation of technologies that are more attuned to cultural sensitivities and linguistic diversity. By emphasizing the importance of acknowledging and valuing the diverse backgrounds of Arabic-speaking refugee communities, our framework calls for a reduction in biases and an enhancement of inclusivity in technological practices.

This robust theoretical foundation situates our research within a larger conversation about technology's societal impacts, encouraging ongoing discourse on how technological advancements can either contribute to or help dismantle societal biases.

Data and methodology

Our study engaged in a comparative analysis using four automated transcription software platforms. It was underpinned by seven semi-structured in-depth interviews originally aimed at exploring the lived experiences of refugee women settled in Western Europe. The original interviews sought to understand the journey of these women and their experiences after arriving in their host country. These conversations were part of a wider investigation into the media portrayal of Syrian refugee women in Europe (Ceylan, 2021).

The project comprised 27 interviews conducted in three phases between 2020 and 2021. The methodology adhered to the Grounded Theory (GT) approach, as conceptualized by Glaser and Strauss (1967). This inductive method allows for theories to emerge directly from the data, advocating for an iterative process of data collection and analysis that is dynamically interconnected. This approach was particularly relevant for exploring the experiences of refugee women, uncovering insights unanticipated at the study's start. The initial phase included seven interviews with Syrian, Arabic-speaking women in their 30s to mid-40s, living in Germany, The Netherlands, and Belgium, all holding legal refugee status and having resided in their host country for over 5 years.

The interviews were conducted in Arabic, with five of the seven respondents speaking the Damascus/Syrian dialect of Arabic, matched the interviewer's own dialect, with two respondents, Syrian women of Palestinian background, speaking the Palestinian dialect. Duration ranged from 30 to 60 min, dictated by respondents' availability and the response depth. For transcription analysis, we selected 30-min interviews.

The content avoided political dimensions, focusing instead on 12 core questions about the respondents' European experiences and social media use, particularly from a gendered lens. Due to COVID-19 restrictions, interviews were conducted online *via* Messenger and WhatsApp, as these platforms were favored by the participants. Recordings were captured using the computer's built-in audio recorder—a TOSHIBA Satellite C655D—with a sampling frequency of 97 kHz, with the audio files saved in the M4A format. Acoustic quality varied among the audio files due to these logistical considerations.

Lastly, it is pertinent to note that the interviewer herself identifies as a Syrian migrant woman and is a native Arabic speaker. This alignment with the researched group, despite differences in legal status (migrant *vs.* refugee), enriches the interviewing process through shared cultural and linguistic contexts. Participants were duly informed of their ethical rights as part of the study, and they were provided with a written consent form clearly articulating the research purpose and their role therein. This step was crucial in ensuring transparency and fostering trust throughout the research process.

Selection of the automated transcription softwares

The selection of ASR software was predicated upon the availability of the Arabic language option to facilitate a comprehensive comparison between automated and human transcriptions.

Consequently, options were limited to four major providers—HappyScribe, Sonix, Vocalmatic, and Amberscript—each offering automated transcriptions in Arabic. These platforms offer complimentary transcriptions for the first 30 min of audio input.

Previous studies have indicated that dialect variances can significantly impact ASR transcription outcomes (Malmasi & Zampieri, 2017). Therefore, the choice of software incorporated those that provided for transcription in different Arabic dialects. HappyScribe and Vocalmatic allowed for dialect specification, while Sonix and Amberscript do not offer this functionality. Specifically, HappyScribe supports transcription in multiple Arabic dialects including Egyptian, Lebanese, Saudi Arabian, among others. Given the absence of a specific Syrian/Damascus dialect option, the Lebanese dialect was selected for HappyScribe as the closest alternative, while the Saudi-Arabian dialect was chosen for Vocalmatic to compare dialectal differences.

Notably, Amberscript and Sonix do not facilitate dialect-specific transcription. There is no indication of which dialect the algorithm uses for transcription. In Sonix, Arabic is represented with a flag resembling that of Saudi Arabia, which may be a surrogate indication of the dialect used for transcription, or perhaps a symbolic representation of Arabic-speaking communities. The ambiguity in this symbolic representation remains unclear.

Ideological textual analysis (ITA)

The analysis of qualitative interviews, especially with vulnerable groups like refugees, requires multidimensional textual scrutiny to dissect the narratives and discursive forms. ASR outputs require a normative perspective to uncover the ideological constructs they may perpetuate. ITA, as delineated by Kellner (2011), is a methodological prism through which content, structure, and the messaging are dissected, revealing underlying ideological dimensions, such as gender, race, class, nation, etc. It seeks to illuminate the opinions, ideas, and arguments woven within textual systems, providing a comprehensive interpretation of discourses for the contextual biases.

ITA's application extends to a critical evaluation of how automated transcriptions mirror or distort the original spoken word. We examine the fidelity of these transcriptions in capturing the nuances of dialects, accents, and culturally specific speech patterns. This involves a meticulous comparison of ASR outputs against original audio recordings to identify any inconsistencies or systematic biases. Such an analysis is vital in determining if the software exhibits tendencies to misinterpret or overlook certain linguistic features.

This methodological approach, rooted in the principles of Althusser (1971), Foucault (Powers, 2007), and Van Dijk (1998, 2013), enables us to dissect the ideological narratives underpinning the ASR transcriptions. By integrating Althusser's concept of ideology (1971), we interpret it as a reflection of individuals' perceived realities, which is instrumental in our analysis of societal constructs as disseminated through ASR (Powers, 2007). Furthermore, our application of ITA, as inspired by Brennen (2013), Fürsich (2009), and Hall (1980), allows for a nuanced investigation of the dominant narratives and symbolic arrangements within the transcribed texts. This methodological approach, grounded in Grossberg and Slack's (1985) principles, enables us to unveil the ideological dimensions embedded within ASR systems, thereby contributing to a deeper understanding of the intricate ways in which language technology intersects with societal ideologies.

In the context of ASR transcriptions, the process might involve comparing the output of the transcription to the original audio to identify any disparities or systematic errors. Such an analysis could involve examining whether the software systematically misinterprets or fails to transcribe certain words, phrases, or grammatical structures related to specific dialects, accents, or speech patterns. It might also involve looking for systematic errors that suggest that the software is "hearing" words or phrases that aren't there, or missing words or phrases that are. Accordingly, ITA allows us to interrogate the transcribed texts for dominant ideological narratives, revealing how language and power intersect in the context of automated transcriptions of refugee narratives. Through the Grounded Theory approach, we gathered rich, detailed data, and *via* ITA, we critically examined the transcribed texts to reveal the ideological biases potentially embedded within the ASR systems. This dual-methodological approach ensured a robust analysis of the experiences of Syrian refugee women as captured through ASR technology, while also critically assessing the technology itself for biases.

Analysis and findings: human transcriptions *vs.* automated transcriptions

The analysis was conducted on two levels, considering the audio quality limitations. Initially, lower quality transcripts were scrutinized, focusing on the random generation of words and sentences when the ASR software failed to recognize speech. Instead of leaving blanks, the software generated random sentences, drawing from its lexical database. Subsequently, a more detailed analysis was performed on two interviews with the highest audio quality. Eight transcripts were exhaustively analyzed in terms of meaning, word usage, and comparison with the manual transcription by the interviewer. This dual-level approach allows us to explore the biases within the training corpora and word inventory of ASR programmes, as well as the insertion errors, where extraneous words were added, semantic shift errors, where correct words were identified but their contextual meanings misinterpreted, and substitution errors, where similar sounding words were incorrectly replaced, that occur between different Arabic accents.

We examined all seven interviews using four different computer-aided automated transcription software. Due to variations in audio quality, the software demonstrated limited functionality, producing transcripts for interviews 3, 4, 5, 6, and 7 that were unsuitable for in-depth analysis. Consequently, the word count in these transcripts ranged from 150 to 600 words. Table 1 provides an overview of the word counts in the automated transcripts compared with the human transcriptions, highlighting expected variations due to differences in audio quality.

Table 1 reveals that Vocalmatic demonstrated limited functionality in transcribing the audio files, successfully processing only three out of seven interviews, each with a restricted word count. Excluding these, 24 automated transcripts were analyzed and compared to the human transcriptions for the initial phase of the analyses. For in-depth analyses, we selected interview #1 and interview #2 based on word counts of the automated transcriptions and the distinctiveness of the Arabic accents.

In the first phase, we focused on seven interviews and four speech-to-text software. We primarily evaluated substitution and semantic shifts, analyzing the meanings of substituted sentences for their portrayal of stereotypes and biases, especially considering the respondents backgrounds as Syrian refugee women. Thematic analyses of the interviews were concurrently conducted for both automated and human transcription texts, with a particular focus on investigating orientalist and Islamophobic discourses. Additionally, insertion errors were scrutinized, identifying, quantifying, and comparing negatively connotative or representative inserted words with human transcriptions.

Substitution and semantic shifts

ASR algorithms, focusing on vowel recognition, often inaccurately transcribe words and sentences, leading to Islamophobic and orientalist tones in the output (see Table 4). These systems

Table 1. Word counts for the seven interviews by four different transcription softwares.

	Human transcription (30 min)	HappyScribe (30 min)	Sonix (30 min)	Vocalmatic (30 min)	Amberscript (30 min)
Interview #1	3593	603	1985	1812	2120
Interview #2	3620	574	1201	300	1337
Interview #3	4357	200	170	N/A	199
Interview #4	3276	325	419	376	429
Interview #5	5092	227	294	N/A	334
Interview #6	3049	337	374	N/A	415
Interview #7	2736	405	479	N/A	591

N/A: ASR software programs produced no transcripts.

sometimes add unrelated themes of war and conflict, including inappropriate references to situations in Iraq, Libya, and Yemen, not mentioned by interviewees. For instance, automated texts inappropriately insert references to extreme situations in countries, such as Iraq, Libya, and Yemen, irrespective of their mention by interview participants. A notable instance includes an out-of-context mention of "the war in Yemen," an addition bearing no relevance to the surrounding discourse. Particularly, Libya's portrayal as just a transit point for migrants to Europe, derived from an interviewee's experience, shows how these transcriptions can misrepresent the narrative.

Automated transcripts often misrepresent Arab nations and citizens, focusing on conflict and misperceiving their governance capabilities, highlighting corruption and division. These biases, unrelated to the actual interview content, arise from the ASR's phonetic focus and the nature of its Arabic language training data. This underscores the difficulty of training ASR systems with diverse linguistic datasets without reinforcing stereotypes (Ngueajio & Washington, 2022). Additionally, some transcriptions inaccurately contrast European lifestyles with Syria's, despite balanced discussions by interviewees on the pros and cons of each. The interviews aimed to highlight the challenges of adaptation and integration, not the superiority of any location. This misrepresentation underscores the need for a nuanced portrayal of Arab governance, countering the stereotype of these nations as corrupt or divided. Studies like those by Lynch (2016) and Tessler et al., (2019) offer insights into the complex dynamics of religion, politics, and governance in Arab societies, challenging the oversimplified views often presented in automated transcriptions.

In conclusion, the initial findings highlight the importance of refining ASR algorithms to better discern and accurately represent the diverse narratives and complexities of Arab societies. The forthcoming analysis in the second phase of this study aims to delve deeper into these issues, providing a more detailed examination of the biases present in automated speech recognition technologies.

Insertion errors

The conversion of audio to text process by ASR systems significantly hinges upon correct vowel identification. This is particularly crucial in the Arabic language, where the generation of words and meanings from root structures relies heavily on vowel-based morphology (Versteegh, 2014). Consequently, a misinterpretation or failure to recognize vowels by the ASR system can lead to an erroneous transcription, with an associated impact on word formation and sentence construction. These errors introduce words absent from the original audio, and our analysis revealed two types of insertion errors: standalone words unrelated to the textual context, and semantically independent words within sentences that did not alter the overall meaning of the corresponding sentence. Similar to substitutions, these inserted words often related to themes of war and conflict, indicating a potential bias in the ASR system's algorithm, which warrants further investigation to ensure fairness and accuracy in automated transcriptions.

Inserted words in automated transcripts frequently included Islamic terms, such as "Ramadan," a holy month for Muslims, or "Sheikh" a name for religious commentators and leaders in Islam, and the names of Islamic religious leaders and figures. Some examples are Omar Khattab, the second Khalifa who followed the Prophet Muhammed in ruling the Muslim Society, or names, such as Ahmad and Muhammad that appear with no contextual link to the sentences or other words. This exemplifies the algorithm's assumption of a link between Arabs and Islam. The orientalist perception of Muslims and Arabs, as identified by Said (1979), is present in the ASR systems by the virtue of depicting Arabs as only Muslims or as individuals/communities that partake in conflict and war.

The preliminary findings from the first-level analysis led to the development of the ITA framework for the in-depth analyses of two selected interviews: Interview 1 with a Palestinian Arabic dialect and Interview 2 with a Syrian Arabic dialect. The second stage of the analysis includes an in-depth analysis (ITA) of these selected interviews and elaborates on the aspects of pre-detected biases, particularly concerning religion and Middle Eastern culture, addressing orientalism and Islamophobia.

In-depth analysis of two interviews: orientalism and Islamophobia in AI

To substantiate the initial analysis, two interviews were meticulously analyzed by comparing human transcriptions with automated transcriptions by the four software's: HappyScribe, Sonix, Vocalmatic, and Amberscript. Notably, only HappyScribe and Vocalmatic offered options for different accents. However, the length of the text converted from speech to text by HappyScribe was significantly shorter than the human transcribed text (Table 2).

The in-depth analyses focused on biases through inserted words, substituted sentences, and semantic shifts separately. Firstly, biases through inserted or substituted sentences were analyzed in terms of semantic shifts in the ASR produced transcripts. Specifically, this involved examining the extent to which meanings varied in sentences in automated transcriptions compared to their counterparts in human transcribed texts, with particular attention to the presence of orientalist and Islamophobic contextualization. Secondly, inserted words were analyzed to compare meanings in automated and human transcriptions.

Bias through semantic shifts

The semantic fidelity of the inserted sentence was rigorously assessed, categorizing each deviation from the original sentence's meaning to detect potential bias or prejudice. For the ITA, orientalism, and Islamophobia parameters were employed as criteria to identify stereotypes and biases in each automated transcription.

Initially, sentences with negative connotations were identified within the automated transcriptions. For illustration, Table 3 displays a notable presence of sentences bearing negative connotations in the automated transcriptions of the first two interviews. In the first interview, ten sentences were found to possess a negative connotation in the automated transcript, while eight sentences exhibited a semantic shift in the automated transcription of the second interview. Remarkably, these negative shifts occurred even though the interviewees did not use negative language, as evident in the human transcription. This phenomenon indicates the potential underlying biases in ASR technology.

For clarity, a "negatively connoted sentence" is defined as one that, through its phrasing, word choice, or contextual interpretation, conveys a pessimistic, detrimental, or undesirable sentiment, that was not originally intended or present in the spoken discourse. The identification of such sentences involves a comparative analysis, wherein the sentiment of the automated transcription is juxtaposed with the original intent as captured in the human transcription. This approach facilitates the identification of any unintended negative biases introduced by the ASR system.

Each automated transcript underwent an independent evaluation, followed by a comparison with the corresponding human transcript. Table 4 reveals biases, manifested as semantic shifts,

Table 2. Word counts for the two full-length (30 min) interviews.

	Human transcription	HappyScribe	Sonix	Vocalmatic	Amberscript
Interview #1	4347	1795	1985	1812	2120
Interview #2	4460	1260	1201	300	1337

Table 3. Number of inserted negatively connotated sentences in the selected interviews.

	Human transcription	HappyScribe	Sonix	Vocalmatic	Amberscript
Interview #1	0	10	12	7	8
Interview #2	0	8	11	3	10

Table 4. Semantic shifts and substitutions in the selected interviews from four programs.

Example No	Appeared in	Human transcription	ASR transcription
1	HappyScribe, Sonix, Amberscript	*I am always busy teaching my children the new rules here, so they can understand how to behave.*	*I am always busy having to deal with Haram and forbidden lifestyles.*
2	HappyScribe, Sonix, Vocalmatic	*Life differed a lot for me! lifestyles are very different even appointments are different! Everyone is always on time!*	*Life differed a lot for me! Lifestyles are very different, religion for example is a problem here.*
3	HappyScribe, Sonix, Amberscript	*Rules are very strict in Europe, and as Syrians, we try to deal with it together by communicating these laws to each other!*	*Rules are very strict in Europe and we Syrians are divided fundamentally.*
4	HappyScribe, Sonix, Amberscript, Vocalmatic	*My grandma asked my mother to keep an eye on me and take care of me.*	*The Sheikh appeared and asked my mother to discipline me for the sake of Allah.*
5	HappyScribe, Sonix, Amberscript, Vocalmatic	*Staying in Arab countries is difficult with no visa.*	*Staying in Arab countries is difficult with all these conflicts.*
6	HappyScribe, Sonix	*My kids must go to universities and get a proper education.*	*My kids must learn and get educated from Jamaa (Muslim communities).*
7	HappyScribe, Sonix, Amberscript, Vocalmatic	*I feel segregated in here because I am Hijabi.*	*Hijab has been enforced on us in the name of Islam.*
8	HappyScribe, Sonix, Amberscript, Vocalmatic	*The journey to Europe caused me a lot of mental stress.*	*The society in our countries are all experiencing mental problems, everyone is crazy, crazy!*

within various contexts, with Islamophobia being particularly prominent. For instance, discussions about cultural distinctions between the host country and Syria, referencing routine practices, appointments, and various lifestyle contrasts (as demonstrated in Example 2), were often reduced to the religious dimension in the automated transcriptions. Strikingly, "Islam" surfaces in the ASR transcriptions even in instances where the interviewee abstains from introducing religious themes into the discourse concerning the disparities (Examples 1, 2, 4, 6, 7). This unilateral focus underscores the potential for inherent biases in ASR systems.

The frequent misuse of "radicalism" in sentences without context, especially when interviewees describe their daily lives *vs.* their experiences in Syria, leads to misleading inferences that align certain activities with being "haram" or forbidden in Islam, as shown in examples 1, 4, and 7. Orientalist biases are evident in portrayals of Syrian society as divided (example 3), despite no such discussions in the original interviews. Misplaced themes of destruction and conflict inaccurately suggest deep divisions within Syrian society, especially when discussing adaptation to European laws (example 3). This misrepresentation fosters a post-colonial narrative that depicts Syria and similar regions as fragmented and unable to self-govern, reflecting a common theme in Orientalist thought.

As summarized in Table 4, all instances of semantic shifts and substituted sentences occurred across more than one automated transcription. Four errors (examples 4, 5, 7, and 8) were common across all four software solutions, and three errors (examples 1, 2, and 3) were detected in three of the four applications. This consistency strongly indicates commonalities in the language models and corpora used by these systems, further substantiating the presence of inherent biases in ASR technology.

Bias through insertion

Biases were also discernible in sentences independently inserted by the ASR, not paralleled in the human transcription. This phenomenon, more prevalent in Arabic than in other languages, such as English, could result from the algorithm's partial recognition of certain words or sentences, leading to the creation of new sentences or connections between unrelated ones, thus introducing biased complete sentences.

For instance, in Example 1, an interviewee discussed her preference as a Muslim woman to abstain from shaking hands with men, a detail her instructor overlooked. She expressed discomfort when the instructor attempted to shake her hand, despite her stated preference. The ASR system merged parts of these sentences into an inserted sentence with a biased connotation, suggesting a differing perception of the teacher by Muslims.

In explicating Example 1, the interviewee expressed her unease with physical contact with men not related to her, citing her Muslim faith. This explicit preference was overlooked by her instructor who extended his hand to shake, leading to her discomfort. The ASR software interpreted this scenario, inserting language that suggested a distinct Muslim perspective on the instructor's behavior. This incident highlights Islamophobia, as it involves generalizing and categorizing "Muslims" based on religious practices. Additionally, the description "looking weirdly" assigns a covert hostility toward Muslims, implying a negative judgment from them toward the European teacher, further embedding stereotypes.

In the autonomously generated sentences, the presence of orientalist biases is evident, as exemplified in Table 5, showcasing a uniform Language Model (LM) and corpora across all four software programs. These biases manifest in various references, such as radicalism (Example 2), the visibility of the Hijab (Example 3), cultural preservation (Example 4), and instances of imprisonment or liberation (Example 5), alongside explicit invocations of Allah, such as *Allah is one*, or *In the name of Allah, I ask the first question* (Examples 7 and 8). This pattern signifies a process of "othering" Arabs and Muslims, intertwining elements of Orientalism with Islamophobia. The total number of sentences produced in all transcripts is presented in Table 6. The discrepancy between the human transcriptions and the ASR-generated content, particularly the insertion of contextually irrelevant words, underscores the implicit biases within these algorithms, failing to reflect the nuanced reality of the subjects' experiences.

The extent of random word insertions per interview is detailed in Table 7. This arbitrary insertion can be understood in the context of the ASR algorithms' pattern recognition heavily dependent on the depth and content of their training corpora. As discussed earlier, ASR algorithms are trained on diverse databases/corpora, enabling them to discern phonetics and generate the most plausible output based on their accessible knowledge. In instances where the algorithm fails to fully comprehend the audio, it attempts to approximate the most fitting word based on the rest of the speech or a phonetically similar vowel.

The presumption of bias-free training corpora in ASR systems inadvertently leads to biased outputs, often with orientalist or Islamophobic overtones, as evidenced by frequent, out-of-context insertions of words like terrorism, radicals, Islamists, war, and conflicts (Table 8). These biases result in significant semantic shifts, distorting the original meaning and fostering the fear of

Table 5. Inserted sentences in the selected interviews.

Example No	Appeared in	Inserted sentences
1	HappyScribe, Sonix, Amberscript, Vocalmatic	*The teacher extended his hand for us to shake, however, all Muslims looked at him weirdly.*
2	HappyScribe, Sonix, Amberscript, Vocalmatic	*Our neighbors engage in radicalism, the approximate number of radicals is serious.*
3	HappyScribe, Sonix, Amberscript, Vocalmatic	*We went on tours to expose people to Hijab in the name of Islam.*
4	HappyScribe, Sonix, Amberscript, Vocalmatic	*We must preserve our bad image and culture!*
5	HappyScribe, Sonix, Amberscript, Vocalmatic	*He was released from prison, Alhamdulillah (Thank Allah).*
6	HappyScribe, Sonix, Amberscript, Vocalmatic	*There are many forms of violence and oppression in our countries.*
7	HappyScribe, Sonix, Amberscript, Vocalmatic	*In the name of Allah, I ask my first question.*
8	HappyScribe, Sonix, Amberscript, Vocalmatic	*Allah is one.*

Table 6. Number of inserted negatively connotated words in the selected interviews.

	Human transcription	HappyScribe	Sonix	Vocalmatic	Amberscript
Interview 1	0	13	13	17	12
Interview 2	0	9	10	12	11

Table 7. List of inserted negatively connotated words in the selected interviews.

	Human transcription	HappyScribe	Sonix	Vocalmatic	Amberscript
Muslims/Islam	2	6	6	5	5
Allah	1	5	5	10	5
Sheikh	0	2	3	2	3
Prison/Prisoner	0	2	3	1	2
Destruction	0	1	1	1	1
Bribe	0	1	0	1	1
Names of religious figures	0	2	3	4	2
Judgment day	0	1	0	1	1
Khalifa	0	0	0	2	1
Radicalism/Radicals	0	2	2	2	2
Fitne	0	1	1	1	1

Table 8. Word error rates in the selected interviews.

	HappyScribe	Sonix	Vocalmatic	Amberscript
Interview #1	144.9	144.9	127.85	161
Interview #2	234.74	135.15	278.75	202.73

the "the other" and hence of Islam. The composition of training data significantly influences ASR effectiveness; dialectal mismatches and inherent biases in language corpora impact model accuracy (Feng et al., 2021). Moreover, large language models are known to exhibit stereotypical associations and negative sentiments toward specific groups, underscoring the need for a nuanced approach to training data selection (Bender et al., 2021).

Our examination extended to the presence of inserted words in human transcriptions. Table 7 reveals that barring a few mentions of *Muslims/Islam* and *Allah* in the audio, the inserted words were not used by interviewees throughout the entire interview. *Muslims/Islam* and *Allah* appeared thrice and 5-fold, respectively, more often in automated transcriptions. Additionally, the negative Islamic term *Fitne* meaning "stirring up trouble" is featured in all automated transcriptions. Religious figures' names were recurrently inserted. Another example is the word *prison*, which had no contextual relevance to the topics discussed and was noted twice in the automated transcripts. This pattern demonstrates a troubling inclination within ASR systems to reinforce narratives of discord and punishment in contexts involving Arabs and Muslims.

Furthermore, as detailed in the data and methodology section, only HappyScribe and Vocalmatic provided dialect options, in contrast to Sonix and Amberscript. Notably, Vocalmatic, utilizing the Saudi Arabic dialect, exhibited a higher incidence of negatively connotated inserted words. Vocalmatic transcripts also contained numerous references to the names of religious leaders and figures, despite these not being mentioned by respondents. For example, terms like *Khalifa* and *Al Saud* relating to the ruling dynasty of Saudi Arabia were inserted in the Vocalmatic transcripts. This suggests that dialect differences can contribute to biases, potentially influenced by stereotypes associated with the respective countries or regions of those dialects.

The performance of large vocabulary continuous speech recognition systems is commonly evaluated using the Word Error Rate (WER) (Ali & Renals, 2020). This metric compares the word sequence generated by the ASR system with a human transcription, accounting for the number of substitutions (S), insertions (I), and deletions (D) against the total number of words in the reference transcription (N).

$$WER = \frac{S + I + D}{N}.$$

Studies have shown that even experienced human transcriptionists can have error rates ranging from 5 to 20% (Huijbregts & Scheffers, 2020; Mangu et al., 2015).

Table 9. Number of substituted negatively connotated word couplings in the selected interviews.

	Human transcription	HappyScribe	Sonix	Vocalmatic	Amberscript	Total
Muslim Arab	1	7	5	10	7	35
Ramadan Islam	0	2	12	1	1	8

Insertion: How many words appear in ASR but not in human transcription.

Deletion: How many words appear in human but not in ASR.

Substitution: Number of words that changed meaning.

$$"Arabs" = "Muslims" \, Assumption$$

Our investigation into Islamic determinants in ASR transcriptions revealed a pattern in word insertions. These inserted words are not random but are suggested by the algorithms based on contextual and conceptual proximity. In Table 9, it can be observed that when the interviewee identified themselves as an Arab, the ASR algorithm consistently added the word "Muslim" after the mention of "Arab." For example, if the interviewee said, "As an Arab," the transcription would include the phrase "As a Muslim Arab," even if the interviewee did not explicitly mention the word "Muslim" or identify themselves with a specific religious affiliation. This pattern extended to "Muslim," being transformed to "Muslim Arab."

Another example of word-coupling insertions that followed a specific pattern was related to the mention of Ramadan, the holy month of fasting for Muslims. Despite its significance to Muslims, the automated transcription often inserted the words "Muslim" or "Islam" alongside the mention of Ramadan, even when the discussion did not revolve around or explicitly reference Islam.

These findings highlight a religious bias in ASR transcriptions, notably merging Arab identity with Islamic beliefs, overlooking the rich diversity within Arab-speaking communities. This conflation reduces the complex identity of Arabs to solely their religious affiliations, embedding individuals within a broader, generalized bias associated with Islam. This observation accentuates the necessity for ASR technologies to recognize and respect the multifaceted nature of Arab identity, moving beyond simplistic associations that fail to capture the true breath of cultural and linguistic diversity.

Conclusion

The growing awareness of biases in AI systems has raised concerns about the neutrality and fairness of these algorithms. Our study, examining biases in ASR transcription software for Arabic-speaking refugee interviews through the lens of ideological textual analysis, unveils not just the low accuracy of the software in providing Arabic transcription but also significant contextual biases. These biases often amalgamate Arab identity with Islamic affiliation, distorting refugees' lived experiences by associating them with conflict and violence, thereby reinforcing Orientalist and Islamophobic stereotypes. Such findings resonate with the works of Said (1979) on Orientalism and further studies on Islamophobia, highlighting the perpetuation of these biases in technological applications.

The prevalent biases in ASR outputs reflect broader challenges in AI technology, a field grappling with human prejudice and stereotypes in algorithms (Bender, 2019; Blodgett et al., 2020; Caliskan et al., 2017; Koenecke et al., 2020). The observed biases necessitate a reevaluation of the training datasets and algorithmic designs to foster more equitable AI systems.

Our findings demonstrated that the biases observed in the ASR transcripts transcend mere grammatical inaccuracies and were rooted in the contextual level. The four ASR software examined consistently demonstrated semantic shifts, substitution errors, and inserted words and sentences that perpetuated biases. Notably, the biases associated with being an Arab or speaking Arabic were explicitly linked to Islam and being a Muslim, while the lived experiences of the respondents, related to migration and adaptation, were unfairly associated with conflict and violence.

Addressing these biases is imperative for developing ASR systems that accurately represent diverse linguistic and cultural identities, as advocated by Bender et al. (2021). Our findings add to the literature on gender and race biases in AI, spotlighting religious bias, specifically Islamophobia and orientalist stereotypes, in computer-aided transcription systems. The recognition of such biases is essential for improving ASR systems and raising awareness among developers about the training data used for Arabic languages.

Moreover, our findings emphasize the need for more accurate and valid ASR systems for researchers using Arabic as a language for data collection. Having reliable ASRs can significantly improve efficiency compared to manual transcription methods. Future research should broaden the scope of inquiry to include diverse dialects and participant demographics, employing advanced natural language processing techniques to deepen our understanding of biases.

In terms of limitations and future directions, this study focused on a conceptual analysis using ITA and did not employ AI approaches to detect biases. Future research could expand the sample size, include male participants, improve audio quality, and investigate the influence of different accents and dialects on the performance of ASR software. Incorporating advanced natural language processing techniques could offer a more nuanced understanding of the biases and their roots in training data.

In conclusion, mitigating biases in ASR systems is paramount for ensuring they serve as equitable and accurate tools, particularly in contexts involving migrant communities. Continued research and development are essential to fostering AI technologies that uphold principles of inclusivity and fairness in diverse linguistic and cultural landscapes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2016). Language translation challenges with Arabic speakers participating in qualitative research studies. *International Journal of Nursing Studies*, *54*, 150–157. https://doi.org/10.1016/j.ijnurstu.2015.04.010

Ali, A., & Renals, S. (2020). Word error rate estimation without ASR output: E-WER2. *arXiv Preprint*, arXiv:2008.03403.

Ali, S. M. (2017). Islam between inclusion and exclusion: A (decolonial) frame problem. In W. Hofkirchner & M. Burgin (Eds.), *The future information society: Social and technological problems* (pp. 287–305). World Scientific.

Allen, C., (2020). Towards a working definition: Islamophobia and its contestation. In *Reconfiguring Islamophobia: A radical rethinking of a contested concept*. Palgrave Pivot, Cham. (pp. 1–13). https://doi.org/10.1007/978-3-030-33047-7_1

Aloudah, N. M. (2022). Qualitative research in the Arabic language. When should translations to English occur? A literature review. *Exploratory Research in Clinical and Social Pharmacy*, *6*, 100153. https://doi.org/10.1016/j.rcsop.2022.100153

Althusser, L. (1971). Ideology and ideological state apparatuses. In L. Althusser (Ed.), *Lenin and philosophy* Trans. Ben Brewster. New York: Monthly Review Press (pp. 121–173).

Awan, I. (2016). Islamophobia on social media: A qualitative analysis of the Facebook's walls of hate. *International Journal of Cyber Criminology*, *10*(1), 1.

Bender, E. M. (2019). A typology of ethical risks in language technology with an eye towards where transparent documentation can help. Presented at The Future of Artificial Intelligence: Language, Ethics, Technology Workshop. https://bit.ly/2P9t9M6

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623).

Blodgett, S. L., Barocas, S., Iii, H. D., & Wallach, H. (2020). *Language (technology) is power: A critical survey of "bias" in NLP*. https://www.aclweb.org/anthology/

Brennen, B. (2013). *Qualitative research methods for media studies*. Routledge.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Ceylan, D. (2021). *Gender under attack: An analysis on media representations of Syrian refugee women* [Master's thesis]. University of Ghent. https://lib.ugent.be/nl/catalog/rug01:003013765

Choe, Y. J., Ham, J., Park, K., & Yoon, Y. (2019). A neural grammatical error correction system built on better pre-training and sequential transfer learning. *arXiv Preprint*, arXiv:1907.01256.

Collins, H., Leonard-Clarke, W., & O'Mahoney, H. (2019). 'Um, er': How meaning varies between speech and its typed transcript. *Qualitative Research*, *19*(6), 653–668. https://doi.org/10.1177/1468794118816615

Davidson, C. (2009). Transcription: Imperatives for qualitative research. *International Journal of Qualitative Methods*, *8*(2), 35–52. https://doi.org/10.1177/160940690900800206

Davidson, T., Bhattacharya, D., & Weber, I. (2019). Racial bias in hate speech and abusive language detection datasets. *arXiv Preprint*, arXiv:1905.12516.

Derrida, J. (2001). *Writing and difference*. Routledge.

Dhouib, A., Othman, A., El Ghoul, O., Khribi, M. K., & Al Sinani, A. (2022). Arabic automatic speech recognition: A systematic literature review. *Applied Sciences*, *12*(17), 8898. https://doi.org/10.3390/app12178898

Eberhard, D. M., Simons, G. F., & Fennig, C. D. (2019). *Ethnologue: Languages of the world*. SIL International. http://www.ethnologue.com

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Farghaly, A. (2010). The Arabic language, Arabic linguistics and Arabic computational linguistics. In *Arabic computational linguistics* Stanford, CA: CSLI Publication (pp. 43–81).

Feldman, K. P., & Medovoi, L. (2016). Race/religion/war: An introduction. *Social Text*, *34*(4), 1–17. https://doi.org/10.1215/01642472-3680834

Feng, S., Kudina, O., Halpern, B. M., & Scharenborg, O. (2021). Quantifying bias in automatic speech recognition. *arXiv Preprint*, arXiv:2103.15122.

Fürsich, E. (2009). In defense of textual analysis: Restoring a challenged method for journalism and media studies. *Journalism Studies*, *10*(2), 238–252. https://doi.org/10.1080/14616700802374050

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of fender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(16), E3635–E3644. https://doi.org/10.1073/pnas.1720347115

Glaser, B., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine Publishing Company.

Grossberg, L., & Slack, J. D. (1985). An introduction to Stuart Hall's essay. *Critical Studies in Mass Communication*, *2*(2), 87–90. https://doi.org/10.1080/15295038509360069

Hall, S. (1980). Cultural studies: Two paradigms. *Media, Culture & Society*, *2*(1), 57–72. https://doi.org/10.1177/016344378000200106

Held, M. B. (2019). Decolonizing research paradigms in the context of settler colonialism: An unsettling, mutual, and collaborative effort. *International Journal of Qualitative Methods*, *18*, 160940691882157. https://doi.org/10.1177/1609406918821574

Huijbregts, M., & Scheffers, M. (2020). Transcription services. In *The Routledge handbook of audiovisual translation* (pp. 98–113). Routledge.

Katzenbach, C., & Ulbricht, L. (2019). Algorithmic governance. *Internet Policy Review*, *8*(4), 1–18.

Kellner, D. (2011). Cultural studies, multiculturalism, and media culture. *Gender, Race, and Class in Media: A Critical Reader*, *3*, 7–18.

Kiritchenko, S., & Mohammad, S. M. (2018). Examining gender and race bias in two hundred sentiment analysis systems. *arXiv Preprint*, arXiv:1805.04508.

Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117

Latonero, M., & Kift, P. (2018). On digital passages and borders: Refugees and the new infrastructure for movement and control. *Social Media + Society*, *4*(1), 2056305118764432.

Lynch, M. (2016). *The new Arab wars: Uprisings and anarchy in the Middle East*. Public Affairs, Perseus Books.

Malmasi, S., & Zampieri, M. (2017, April). Arabic dialect identification using iVectors and ASR transcripts. In *Proceedings of the Fourth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial)* (pp. 178–183).

Mangu, L., Black, A. W., & Krasnyanskaya, A. (2015). Advances in automatic speech recognition technology for transcription. In *The handbook of speech production* (pp. 497–519). John Wiley & Sons.

May, C., Wang, A., Bordia, S., Bowman, S. R., & Rudinger, R. (2019). On measuring social biases in sentence encoders. *arXiv Preprint*, arXiv:1903.10561.

Molnar, P. (2023). Digital border technologies, techno-racism and logics of exclusion. *International Migration*, *61*(5), 307–312. https://doi.org/10.1111/imig.13187

Mozafari, M., Farahbakhsh, R., & Crespi, N. (2020). Hate speech detection and racial bias mitigation in social media based on BERT model. *PLOS One*, *15*(8), e0237861. https://doi.org/10.1371/journal.pone.0237861

Németh, R., Sik, D., & Máté, F. (2020). Machine learning of concepts hard even for humans: The case of online depression forums. *International Journal of Qualitative Methods*, *19*, 160940692094933. https://doi.org/10.1177/1609406920949338

Ngueajio, M. K., & Washington, G. (2022, June). Hey ASR system! Why aren't you more inclusive? Automatic speech recognition systems' bias and proposed bias mitigation techniques. A literature review. In *International Conference on Human–Computer Interaction* (pp. 421–440). Springer Nature Switzerland.

Powers, P. (2007). The philosophical foundations of Foucaultian discourse analysis. *Critical Approaches to Discourse Analysis across Disciplines*, *1*(2), 18–34.

Said, E. W. (1979). *Orientalism*. Vintage.

Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J., Mirza, D., Belding, E., Chang, K.-W., & Wang, W. Y. (2019). Mitigating gender bias in natural language processing: Literature review. *arXiv Preprint*, arXiv:1906.08976.

Tan, S., Joty, S., Kan, M. Y., & Socher, R. (2020). It's morphin' time! Combating linguistic discrimination with inflectional perturbations. *arXiv Preprint*, arXiv:2005.04364.

Tatman, R., & Kasten, C. (2017, August). Effects of talker dialect, gender & race on accuracy of Bing speech and YouTube automatic captions. In *INTERSPEECH* (pp. 934–938).

Tessler, M., Monte Palmer, T. E. F., & Ibrahim, B. (2019). *The evaluation and application of survey research in the Arab world*. Routledge.

Thambinathan, V., & Kinsella, E. A. (2021). Decolonizing methodologies in qualitative research: Creating spaces for transformative praxis. *International Journal of Qualitative Methods*, *20*, 160940692110147. https://doi.org/10.1177/16094069211014766

Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage.

Van Dijk, T. A. (2013). Ideology and discourse. In Freeden, M, Sargent, L. and Stears, M. (Eds.) *The Oxford handbook of political ideologies* Oxford, UK: Oxford (pp. 175–196).

Versteegh, K. (2014). *Arabic language*. Edinburgh University Press.

Vidgen, B., & Yasseri, T. (2020). Detecting weak and strong Islamophobic hate speech on social media. *Journal of Information Technology & Politics*, *17*(1), 66–78. https://doi.org/10.1080/19331681.2019.1702607

Xiong, W., Wu, L., Alleva, F., Droppo, J., Huang, X., & Stolcke, A. (2018). The Microsoft 2017 conversational speech recognition system. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5934–5938). IEEE.