Thesis for The Degree of Licentiate of Philosophy

Robot Voices Matter More Than You Think

From Gendering to Inclusion with Gender-Ambiguous Voices

Martina De Cet

Department of Computer Science and Engineering
Chalmers University of Technology | University of Gothenburg
Gothenburg, Sweden, 2026

Robot Voices Matter More Than You Think
From Gendering to Inclusion with Gender-Ambiguous Voices

Martina De Cet

ISSN 1652-876X

Department of Computer Science and Engineering
Division of Interaction Design and Software Engineering
Chalmers University of Technology | University of Gothenburg
Sweden
Phone: +46(0)31 772 1000

Printed by Chalmers Digitaltryck,
Gothenburg, Sweden 2026.

Robot Voices Matter More Than You Think

From Gendering to Inclusion with Gender-Ambiguous Voices

Martina De Cet

Department of Computer Science and Engineering
Chalmers University of Technology | University of Gothenburg

Abstract

Voice shapes social perception quickly and persistently. Beyond transmitting content, it shapes impressions of gender, age, origin, warmth, competence, and authority, which in turn affect how speakers and technologies are judged. In Human–Robot Interaction (HRI) and Human–Computer Interaction (HCI), robots and voice assistants commonly use clearly masculine or feminine synthetic voices. This practice reflects a broader tendency to rely on binary gender cues, both as a general design convention and as a means of making technology appear natural and acceptable. However, these design defaults can reinforce gender stereotypes and role expectations, contribute to masculine-by-default perceptions of technology, and make interactions less welcoming for users who do not align with binary categories. This thesis examines gender-ambiguous synthetic voices as a design intervention that resists clear categorisation within masculine or feminine norms, investigating how such voices are perceived, whether they support inclusion, and whether they influence robot gendering.

This thesis comprises four papers. Paper I presents a systematic review of voices beyond the binary in HCI, revealing inconsistent terminology, limited transparency about voice availability, and a lack of methodological consistency in evaluation. Paper II reports a large-scale perception study evaluating ambiguous voices across trustworthiness, comfort, appeal, anthropomorphism, and aversion, showing that these voices are not experienced uniformly and can elicit more critical responses from nonbinary listeners. Paper III and Paper IV examine how voice influences imagined and constructed robot form using sketching and physical prototyping as methods. Together, these studies show that gender-ambiguous voices can reduce explicit gender attribution and encourage less clearly gendered embodiments, while a masculine-by-default tendency remains.

Overall, this thesis shows that gender-ambiguous voices can soften binary gendering in interaction and robot design, but they do not provide identity-based inclusion without community-informed approaches and careful framing.

Keywords

Human-Robot Interaction, Human-Computer Interaction, Artificial Voices, Gender-Ambiguous Voice, Robot Gendering, Inclusivity, Sketching, Prototyping

List of Publications

Appended Publications

This thesis is based on the following publications:

[**Paper I**] **M. De Cet**, M. Obaid, I. Torre, (April 2025), *Breaking the Binary: A Systematic Review of Gender-Ambiguous Voices in Human-Computer Interaction*
*Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (**CHI**).*
https://doi.org/10.1145/3706598.3713608

[**Paper II**] **M. De Cet**, K. Seaborn, M. Obaid, I. Torre (July 2025), *Hearing Ambiguity: Exploring Beyond-Gender Impressions of Artificial Ambiguous Voices*
*Proceedings of the 7th ACM Conference on Conversational User Interfaces (**CUI**).*
https://doi.org/10.1145/3719160.3736622

[**Paper III**] **M. De Cet**, M. Sturdee, M. Obaid, I. Torre, (March 2025), *Sketching Robots: Exploring the Influence of Gender-Ambiguous Voices on Robot Perception*
*Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction (**HRI**).*
https://dl.acm.org/doi/abs/10.5555/3721488.3721505

[**Paper IV**] **M. De Cet**, N. Hashmati, M. Obaid, I. Torre, (Accepted), *From Voice to Form: How Gender-Ambiguous Voices Shape Physical Robot Design*
*Proceedings of the 2026 ACM/IEEE International Conference on Human-Robot Interaction (**HRI**).*

Other Publications

The following publications were produced during my PhD studies. However, they are not included in this thesis because they fall outside the thesis scope.

[**a**] C. Wang, A. Wikström, L. Pettersson, A. Sarker, **M. De Cet**, G. Diapoulis, M. Obaid, I. Torre (July 2025), *SitBot: A Posture-Mimicking Robot to Reduce Slouching*
International Conference on Social Robotics (ICSR).
https://doi.org/10.1007/978-981-95-2379-5_58

[**b**] N. Walczak, **M. De Cet**, M. Obaid, A. Romanowski (September 2025), *Co-designing for Pain Relief: Recommendations for Assistive Robotic Design*
Human-Computer Interaction – INTERACT 2025 (INTERACT).
https://doi.org/10.1007/978-3-032-04999-5_4

[**c**] **M. De Cet**, M. Cvajner, I. Torre, M. Obaid (July 2024), *Do Your Expectations Match? A Mixed-Methods Study on the Association Between a Robot's Voice and Appearance*
Proceedings of the 6th ACM Conference on Conversational User Interfaces (CUI).
https://dl.acm.org/doi/10.1145/3640794.3665551

[**d**] N. Walczak, **M. De Cet**, A. Romanowski, M. Obaid (December 2023), *From Pain to Design Recommendations for Assistive Robotics*
Proceedings of the 11th International Conference on Human-Agent Interaction (HAI).
https://dl.acm.org/doi/10.1145/3623809.3623950

Acknowledgment

Writing the acknowledgements is probably the most difficult part of this thesis. Trying to squeeze all the love and admiration I have for the people around me into just a few words feels almost impossible and inevitably reductive. Still, I will do my best.

First and foremost, my biggest thanks – truly the biggest, because without you this thesis would not exist – go to my amazing partner, Klara. Your support has been extraordinary, to the point that I sometimes wonder whether it would only be fair to list you as a co-author on my papers. Thank you for your endless patience, infinite love, support, and care.

My deepest thanks also go to my parents, my sister, Giovanni, and my uncle, for being constant and unwavering supporters. Being far away is not easy, but it helps to know that, whenever I think of home, I have something to be proud of and people to share it with. Perhaps in another life, I should have done a PhD on teleportation, just to find a way to come home more often and hug you all whenever I need to (very often). And to you, nonna: I wish I could tell you all about my life now.

I am extremely grateful to my supervisor, Ilaria. Over the past two years, thanks to your guidance and support, I have gained an enormous amount of knowledge and, more importantly, confidence in myself – something I have always struggled with. Your trust in my abilities has meant more to me than I can easily put into words. My other big thanks go to my co-supervisor, Mohammad. Without the passion for research that I gained through my internship and Master's thesis with you, I would never be writing these pages. Thank you for being the first professor who truly saw something in me.

To all my colleagues at Interaction Design and Software Engineering, thank you for the welcoming environment and for the encouragement you have always given me. I am deeply grateful for your support, help, and for the time you dedicated to my research. A special thanks goes to everyone at Interaction Design, who have been kind to me since my time as an intern and continue to be so now that I am your colleague and friend. I truly appreciate you all. A big thank you to Cecilia and Clara for their invaluable support.

A special thanks go to my colleagues and friends. Thank you to Sabi for sharing both the pain and the joy of this journey with me. To Mafalda, for the warmth, love, and a bit of craziness; to Sam, for being such a kind and supportive presence; to Razan, for being such a great office mate; and to Natasha, for being both a wonderful office mate and a good friend. Thank you to Francisco, for making Chalmers a welcoming place for queer people, and for making it so for me as well – something I had not experienced before. Thank you as well to Sjoerd for the academic support over the past months. And lastly, thank you to my dear friend and colleague, Negin. Thanks for being there on the difficult days, and for making everything feel lighter, even when it is just through a bit (or a lot) of gossip at work.

Thank you to all the friends who made, and continue to make, Gothenburg a better place. An enormous thank you to my queer group: because of you, Gothenburg feels like a safer, warmer, and more joyful place. Thank you, Sunny, for the kindness, support, and shared moments that made Gothenburg feel more like home from the day I met you.

Thanks to my choir friends in Trento, who keep my plan B alive in case this PhD does not work out (opening a bakery or a cooking business). Thank you to Giulia, Ofi, and Tere – the friends I miss the most all year round. I wish we could all live in the same country forever.

Finally, thank you to my Swedish family, Kerstin and Anders, for welcoming me into your home from day one, and for being so inclusive, caring, and loving towards me.

Gothenburg, January 2026

Contents

Part I

Summary

Chapter 1

Introduction

What grabs your attention when you meet someone? For me, it is usually the voice. Sometimes it is because the voice is soft and calm. Other times, it is because it has a darker pitch that reminds me of the narrator from the Puss in Boots cassette I loved as a child. And sometimes it is simply because I love the accent of the person speaking. For example, after three years living in Sweden, I can say I genuinely enjoy a thick Swedish accent when a Swede speaks English. People joke that it sounds like listening to Stitch, from Lilo&Stitch, which honestly feels very accurate. But every time I say this to Swedes, they look horrified and say: *"Oh no, Swedish English sounds awful"*.

Honestly, I get it. I feel the same when someone tells me they love an Italian speaking English with a very thick Italian accent. When I ask why, they say it is musical, elegant, and full of passion. I always think, *"Is it?"* I never quite see it that way. But reactions like these make it impossible to ignore how much importance we give to voice and all the small details that come with it: pitch, accent, speed, and many more.

All of this to say that voice matters. It shapes how we see and hear each other. It is everywhere. It is how many people express themselves, how they are understood, and how they are categorised. Yet voice is more than sound. It is not only a biological or a communication tool but also a social, cultural, and political phenomenon [1]. How someone speaks can signal origin or social position, and whose voice is listened to often reflects existing power relations.

Voice carries meaning. People interpret voices through assumptions about gender, race, class, and identity. Listening to a voice involves more than processing auditory cues; we *hear* a person. Rapid judgments then emerge: *Is this a man or a woman? Young or old? Trustworthy or not?* Voice is also political. Public systems often use feminine voices for care-related roles, while masculine voices are usually used for authority and guidance. This pattern aligns with research showing that synthetic male voices are commonly judged as more persuasive and authoritative than synthetic female voices [2]. As you can see, then, voice is never a neutral design choice; it is a site where social norms are repeated, challenged, and negotiated [1].

Not only does voice carry meaning, but it is also central in everyday technology. Voice assistants such as Siri and Alexa are present in hundreds of millions of devices, with more than 320 million units as of 2024 in use worldwide [3]. These systems typically adopt recognisably masculine or feminine voices, and this choice has long been discussed within Human–Computer Interaction (HCI) and Human–Robot Interaction (HRI). Research shows that using a gendered voice shapes how users perceive an agent, affecting trust, likability, and acceptance [4, 5, 6, 7, 8]. Voice conveys such impressions very quickly to listeners, and thus voice design choices quickly shape everyday interactions with technology. While such reasoning may appear intuitive (and harmless), it carries consequences. Gendered voice design can reinforce stereotypes and exclude people whose identities fall outside binary expectations, motivating the need for more inclusive alternatives [4, 9]. Importantly, this binary is reinforced not only through design conventions but also through technical infrastructure: widely used speech corpora largely contain only cisgender[1] male and female voices and implicitly assume normative gender identities [11, 12].

For these reasons, recent HRI and HCI research has started exploring gender-ambiguous (also referred to as gender-neutral or nonbinary) voices to move away from gendered voice design and support more inclusive interactions. In this context, ambiguity refers to vocal cues that do not clearly map onto either masculine or feminine norms. Research in this area is only beginning to develop, with only a few studies published [9, 13, 14], and its broader potential and perception are largely unknown. For this reason, I started my PhD research with the following research question:

> **RQ1**: How do people perceive a gender-ambiguous voice?

Building on this, I then asked whether gender-ambiguous voices, often positioned as a more inclusive option, are perceived differently by listeners of different gender identities. This question is particularly relevant given that nonbinary and gender-diverse people are both strongly affected by binary voice design and remain under-represented in HCI and HRI research, as shown in recent meta-reviews [15, 16]. Therefore, I extend RQ1 with the following:

> **RQ1a**: Do perceptions differ between men, women, and nonbinary participants?

If perceptions and gender attributions differ across gender groups, then voice design choices can have uneven consequences, including who feels addressed, represented, or excluded. As argued by Spiel et al. [17, 18], technological systems always encode assumptions about gender, and designs presented as neutral often reproduce binary norms that privilege cisgender and binary users.

[1] Cisgender refers to a person whose gender identity matches the sex they were assigned at birth [10].

This raises a practical question for voice design: if a system uses a gender-ambiguous voice instead of a clearly gendered one, does it increase inclusivity for users with diverse gender identities? For this reason, I investigate:

> **RQ2**: Can a gender-ambiguous voice increase inclusivity?

Inclusivity, however, is not only about how users feel included or are represented; it is also shaped by how strongly an agent is gendered through design cues. In robotic systems, voice is one of the most salient cues shaping gender attribution and associated expectations. Torre et al. [14] show that a gender-ambiguous voice can reduce the gendering of a robot with stereotypically feminine or masculine traits. Yet evidence in HRI remains limited, and we still lack a clear understanding of when and how ambiguous voices influence robot gendering. For this reason, my third research question asks:

> **RQ3**: Does a gender-ambiguous voice influence robot gendering?

Before concluding this introduction and going more into the details of previous work, I want to address two points: clarifying the terminology used for non-gendered voices to keep the thesis consistent (see 1.1), and outlining how my interest in HRI and this research started (see 1.2).

1.1 Terminology: A Clarification

Across the literature analysed for my systematic review [19], I found out that researchers use several labels for voices that do not align with conventional masculine or feminine categories. Common terms include “gender-ambiguous”, “gender-neutral”, “genderless”, “non-binary”, “gender-free”, and “androgynous”. Although often treated as interchangeable, they refer to different concepts. I define them here as I understand them, after having read works such as [7, 14, 20, 21, 22, 23, 24]:

- **Gender-ambiguous**, **androgynous**: Terms to describe a voice that does not clearly fit into traditional male or female categories by blending both masculine and feminine characteristics, making it difficult to identify the speaker’s gender based solely on vocal characteristics.
- **Gender-neutral**, **genderless**, **gender-free**: Terms to describe a voice that aims to avoid masculine or feminine characteristics, with no alignment to male or female vocal traits, aiming for a balance that does not lean towards either gender.
- **Nonbinary**: A term to describe a voice that incorporates elements from either feminine and/or masculine voices or neither, and is more linked to someone’s identity rather than their voice.

In this thesis, I use “gender-ambiguous” as the term to describe non-gendered voices. Labels such as “gender-neutral”, “genderless”, and “gender-free”, in my opinion, imply a full removal of gendered cues, which is not feasible in practice with technology [22]. Listeners tend to classify voices using binary expectations, and both acoustic features and linguistic content shape these perceptions [7, 25]. "Gender-ambiguous" instead recognises that these voices do not eliminate gendered traits but resist being placed cleanly into one category. It also avoids the appearance-related associations that "androgynous" can carry. Finally, because my PhD initially focused on reducing gendering rather than representing specific identities, I avoided the label "nonbinary", as this primarily refers to a gender identity, not an acoustic category.

1.2 My Interest in HRI: Where It All Began

My interest in HRI began with the research I conducted for my master's thesis, which led me to the questions explored in this work and to my first publication in this PhD journey [26]. In that study, I asked Swedish and Italian participants to either look at a set of robot images and match each one to a voice, or listen to a set of voices and choose the corresponding robot image. Afterwards, I interviewed them about which voice they would prefer a robot to have, with a particular focus on how gender and the context of interaction shaped their choices.

Running that experiment made me realise two things: (1) I wanted to pursue a PhD, and (2) I wanted it to focus on voices. The first became clear because I enjoyed designing, running, and analysing the study, as well as the lab environment. The work was varied, and the mix of tasks made the whole journey engaging. The second point came from seeing how differently Italian and Swedish participants responded. Swedish participants said they disliked female voices for cleaning robots because they reinforce gender stereotypes, whereas Italian participants described female voices as kinder and more pleasant for assisting and support roles. This contrast made me realise how complex and culturally dependent voice perception is. Some Swedish participants also said they would prefer a gender-neutral voice for their robot. My first reaction was *"Wait, what? Gender neutral? What do you mean?"* But once I started looking into it, I realised how interesting and important the idea was and how little I had considered it before.

After thinking more about it, something clicked. I started reading and listening to available non-gendered voices, such as Q, the first genderless voice[2], and I quickly realised this was also the type of voice I would have preferred in a voice assistant. For example, very masculine voices on Google Maps make me feel like I am being mansplained. Very feminine voices, like the one I had on my Alexa, feel strange because they sound as if I am being helped by a woman, which in my head reinforces the stereotype of women being helpful. For this reason, the more I read, the more I thought: *"These non-gendered voices make a lot of sense"*.

[2]www.youtube.com/watch?v=lvv6zYOQqm0

So here I am, presenting my research on gender-ambiguous voices, which connects directly to the idea of gender-neutral voices raised by Swedish participants in that first study. In the next chapter, which serves as related work, I situate this thesis within the growing presence of robots in everyday life and the challenges that follow, focusing on robot gendering and the persistent masculine-by-default effect. I then introduce key perspectives from gender and queer studies to frame why moving beyond binary categories matters, before introducing design strategies that aim to reduce gender attribution. Finally, I examine voice as a central cue in robot perception and summarise the state of the art on voices beyond the binary, highlighting what is known and where gaps remain that motivate my research questions.

Chapter 2

Background

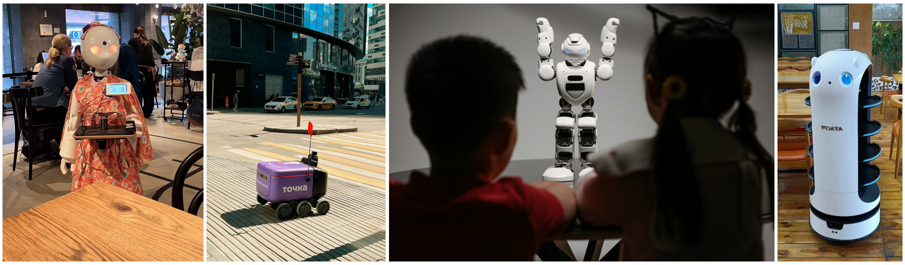

Figure 2.1: Examples of robots used in everyday environments.

Modern society increasingly relies on robots to perform a wide range of tasks, including education, transportation, healthcare, entertainment, and other domains. In the past, robots were primarily used in controlled settings to carry out repetitive tasks based on predefined instructions. Today, robots are increasingly involved in complex and less regulated activities that require direct or indirect interaction with people.

As illustrated in Figure 2.1, robots are already present in everyday environments. The first and last images show service robots currently used in cafés and restaurants, where they take orders and deliver food. Such robots are now common in several countries, including Japan. The first image, for example, was taken by me at the Dawn Café in Tokyo, an avatar-robot café where remote staff "pilots" robots in real time to greet and talk with customers, guide them through the experience, take orders, and perform café service and delivery-style tasks. The second image from the left shows a delivery robot, part of a service that has become popular in some areas of the United States. Instead of relying on human couriers, packages are placed inside the robot, which then autonomously travels to its destination.

Robots are also increasingly present in educational contexts. The second image from the right depicts children interacting with a robot, reflecting a growing interest in social robots as learning companions. Research suggests that social robots can support education by acting as tutors, assisting with repetitive tasks such as attendance, or encouraging peer interaction [27]. For instance, Alemi et al. [28] investigated the use of a humanoid robot (NAO) in English language teaching in Iranian junior high schools. Students who learned with the support of the robot in addition to a human instructor outperformed those taught using traditional methods alone, both in vocabulary retention and learning speed. This study illustrates how robots are not only envisioned for future applications but are already being explored in applied contexts.

Beyond education, another area where robots are highly used is healthcare. In these contexts, social robots can provide companionship, assistance, and emotional support. A growing body of research highlights their potential benefits. For example, Rakhymbayeva et al. [29] used a social robot in long-term therapy sessions with children diagnosed with autism and, in some cases, co-occurring ADHD. The robot supported therapy by offering a range of interactive activities that were adapted over time to each child's preferences and responses, helping to sustain engagement across multiple sessions. Their results showed that children remained attentive and involved over repeated interactions, particularly when sessions included familiar activities tailored to individual needs, highlighting the potential of social robots to support personalised and scalable therapeutic interventions. Recent work [30] also suggests potential benefits for mental wellbeing. This study compared human-, robot-, and digitally guided meditation sessions and found that robot-led sessions produced the greatest reductions in anxiety, despite lower reported trust, highlighting the potential of social robots to provide scalable mental health support.

Taken together, these examples demonstrate that robots are already embedded in many areas of society and that their presence is likely to continue to grow. Alongside this increased adoption, familiar social perception processes become more consequential. For example, people routinely anthropomorphise non-human entities, applying human categories to make sense of them. When applied to robots, this tendency is not inherently problematic, but it can create specific challenges for interaction design and inclusion. One such phenomenon is robot gendering, that is, the tendency for people to assign a gender to robots. Gendering can arise for various reasons, including linguistic conventions, cultural norms, stereotypes about certain roles, and broader associations of technology with masculinity. For example, robots may be perceived as male by default due to grammatical gender (a property of some languages in which nouns are assigned a gender category) or the masculinisation of technology, or as female when associated with caregiving, service, or educational roles.

Such attributions raise important concerns. Indeed, gendered robots can reinforce existing stereotypes and may influence how people behave toward both robots and other humans. For instance, repeatedly interacting with feminine robots in subordinate or service roles risks normalising similar expectations toward women. For this reason, a substantial body of research in HRI has begun to critically examine gender and gendering in robotic systems. The next section explores this literature in more detail.

2.1 Gendering Robots

Gender is a fundamental social category that shapes expectations, behaviours, and social roles in everyday life [31, 32]. When meeting others, people quickly, often unconsciously, use cues such as appearance, voice, name, body shape, clothing, and non-verbal behaviour to judge someone's gender [33, 34, 35, 36]. These processes of gender attribution are also applied to technology. As robots increasingly enter social environments, users draw on the same perceptual habits and gendered expectations they use with humans, interpreting robots through familiar social categories.

Robot gendering refers to the attribution of gender to a robot through design features such as voice, name, appearance, or behaviour [37]. This phenomenon is not merely anecdotal: large-scale analyses of humanoid robot images show that gender attribution is systematic [34]. Most robots are perceived as masculine or gender-neutral, while feminine and androgynous robots are comparatively rare, pointing to an underlying representational bias in robot design.

Gender has been widely used as an experimental variable in HRI research, with a growing body of work examining how robot gender influences user perceptions, attitudes, and behaviour. Early empirical studies demonstrated that even minimal gender cues can activate gendered expectations. For example, Siegel et al. [38] investigated how robot gender and user gender affect persuasion in a museum setting. Participants interacted with a humanoid robot that asked for a donation, with the robot's gender manipulated across conditions. Their results showed complex interactions between robot gender, participant gender, and social context. While women showed little preference, men were more likely to donate to a female robot, and participants generally rated opposite-gender robots as more convincing, dependable, and interesting. These findings were surprising in light of social psychology research suggesting a tendency toward same-gender preference and in-group bias [39].

While Siegel et al. [38] observed a preference for opposite-gender robots, Eyssel et al. [40] found different results. In their experiment, they examined how robot gender and voice type influence user perceptions, emotional closeness, and anthropomorphism. They found that participants evaluated robots more positively and felt a stronger emotional connection when the robot was perceived as sharing their gender. This effect was particularly noticeable when the robot used a human-like voice, suggesting that similarity can facilitate projection and identification processes in HRI.

A robot's role or occupation can also lead to gender attribution. For instance, Aşkın et al. [41] showed that stereotypically gendered tasks led participants to assign gender to an otherwise gender-neutral robot. This suggests that gender attribution does not rely solely on perceptual cues, but is also shaped by contextual and cultural expectations associated with specific roles [42]. Several studies report that aligning a robot's perceived gender with stereotypical role expectations can lead to more positive attitudes, higher trust, and greater acceptance [4, 5]. Tay et al. [43], for example, found that users responded more positively when a robot's gender matched the stereotype of its role, such as feminine robots in healthcare or masculine robots in security roles. Similarly, Fridin and Belokopytov [44] justified the use of a feminine voice for their robot by noting that it operated in kindergarten settings, where most educators are women.

These findings raise an important point. On the one hand, gender cues can improve usability and acceptance by meeting user expectations. On the other hand, they risk reinforcing existing stereotypes. A comprehensive scoping review of gendering humanoid robots [45], shows that while robot gender does not consistently affect likeability or acceptance, it has a robust and recurring impact on stereotyping. In addition, while robots can be gendered in multiple ways, one pattern appears particularly persistent: the tendency to perceive robots as masculine unless explicit cues suggest otherwise. This so-called "masculine-by-default" effect operates across language, culture, and technology, shaping what is perceived as neutral, authoritative, or credible. The following subsection examines how this default emerges, how it has been theorised in feminist scholarship, and why it continues to influence perceptions of robots and other technologies today.

2.1.1 Masculine by Default

For those who may not be familiar, the "masculine-by-default" effect describes the tendency to interpret or represent things through a masculine lens, assuming maleness as the default category. This bias extends beyond science into language and everyday life. For example, in my mother tongue, Italian, which is a gendered language, the grammatical default is masculine. When referring to professional titles such as judge, doctor, or director, the masculine form is typically used. I began reflecting on this (problematic) default when I once designed a poster for my choir and wrote "direttrice" (the feminine of "director") next to my mother's name. She asked me to change it to "direttore" (masculine of "direttrice"), explaining that "direttrice" sounded less serious and less professional. That moment made me realise how deeply the masculine default operates and how it shapes what we perceive as neutral, authoritative, or credible.

When it comes to technology, this tendency becomes even more pronounced. The "masculine-by-default" effect often describes how technologies are perceived as male unless they are explicitly marked otherwise. This idea is not new and has been examined across several disciplines, including feminist theory.

Feminist scholars such as Donna Haraway have shown that what is often called neutral or objective is in fact shaped by masculine norms. Haraway's concept of "the modest witness" [46] helps explain this. She describes how early modern science built its authority on the figure of a detached and rational observer who claimed neutrality but was implicitly imagined as male. This supposedly objective stance became a model for how knowledge and truth were produced, one that excluded embodied, emotional, or situated ways of knowing, often associated with femininity.

This effect remains evident today. Indeed, technology is often treated as masculine by default, including within HCI and HRI research. For example, Han et al. [47] show that when chatbots lack explicit gender cues, users overwhelmingly perceive them as male. This finding illustrates how claims of technological "neutrality" frequently reproduce cultural associations between masculinity, authority, and competence. The fact that this work was published in 2025 suggests that the masculine-by-default effect is not a historical artefact, but an ongoing issue.

Several factors contribute to this pattern. In the case of robots, masculine defaults have been linked to assumptions embedded in technological design [34, 48] and the influence of the "male gaze" in shaping innovation [49], where masculine perspectives implicitly guide what is designed, how it is designed, and which users are prioritised. One reason the masculine-by-default pattern also persists in robotics is that many people learn what a "robot" is supposed to be through media representations long before they encounter real robots. Popular film and animation repeatedly depict robots through character types and design conventions that are readily read through masculine-coded conventions, particularly in roles associated with authority, protection, combat, or technical competence.

Figure 2.2 illustrates how popular film and animation contribute to the masculine-by-default understanding of robots. Widely recognised robotic characters are predominantly male-coded, either through names, voices, pronouns, or narrative roles associated with strength, authority, combat, or technical expertise, as exemplified by characters such as the Terminator T-800, Gundam units, C-3PO, Bender, and Baymax. These representations reinforce the idea that masculinity functions as the unmarked and natural category for robots. By contrast, when female-coded robots appear, such as Rosie or Dot Matrix, they are typically assigned supportive, domestic, or service-oriented roles. This asymmetry suggests that femininity in robotic representation is not absent but role-dependent, becoming visible primarily when it aligns with traditional gendered expectations of care or service.

Addressing this bias requires intentional design strategies that challenge and disrupt these defaults. Incorporating diverse design choices can support less gendered perception and interactions. In a context where binary gender divisions increasingly fail to reflect lived experience, exploring design approaches that move beyond the binary becomes both timely and necessary. To provide a theoretical grounding for this perspective, the next subsection introduces key concepts from gender and queer studies that help frame why and how moving beyond binary thinking is important in HRI (and not only).

Figure 2.2: Examples of robots from popular film and animation (Gundam mobile suit from Mobile Suit Gundam; C-3PO from Star Wars; Bender from Futurama; Rosie from The Jetsons; Dot Matrix from Spaceballs; Terminator T-800 from The Terminator; and Baymax from Big Hero 6).

2.1.2 Gender as a Spectrum: Beyond the Binary

After discussing robot gendering, its effects on users, and the concerns it raises, I believe it is helpful to step back and reflect on the gender framework that often underlies these discussions. In particular, I argue that embracing a "beyond the binary" perspective is necessary, rather than thinking of technology (in this case, robots) strictly as male or female.

Indeed, the world is full of diverse individuals, each with a unique gender identity, and many people do not fit neatly into the "traditional" male/female categories. Even when we examine biology, the "classic" two-sex division is not fully accurate. While most humans can be categorised as male or female, a portion of the population, for example, is intersex[1], highlighting that sex (and gender) cannot be understood solely through biological terms.

Beyond biology, gender identity itself is not limited to two categories. Nonbinary people identify outside the male-female binary, demonstrating that not all individuals experience gender in binary terms. Recognising both intersex and nonbinary people illustrates why restricting our understanding of sex and gender to two categories is insufficient. Instead, it points to sex and gender as existing along a spectrum, and embracing this diversity allows for a more inclusive perspective that acknowledges individuals' varied and lived experiences.

This move toward thinking beyond the binary is not new. Feminist scholars have long pointed out that what we call "sex" and "gender" is never just a biological fact but always shaped by culture and power. Their work shows us that the categories we take for granted, male and female, nature and nurture, are themselves unstable [21, 51]. Gayle Rubin [21] distinguishes anatomical sex, defined by the biological characteristics of bodies, from gender, which encompasses the cultural constructs and norms associated with those characteristics. Rubin's work highlights that traits and roles traditionally linked to men and women are not innate but are shaped by societal and cultural influences.

Building on this perspective, Butler argues in *Gender Trouble* [20] that gender is not something one "has" but something one "does". Gender is performative: it is created through repeated actions such as speaking, dressing, and moving, suggesting that societal norms dictate and reinforce what is perceived as gendered behaviour. However, Butler also emphasises that social norms can be disrupted. When gendered actions are repeated playfully or exaggeratedly, they reveal gender as socially constructed rather than fixed [52]. Drag exemplifies this process: by exaggerating masculinity or femininity, it reveals that all gender expressions are performances, not expressions of some inner truth.

[1] Intersex is an umbrella term for people born with reproductive or sexual anatomy that does not fit typical definitions of male or female. This can include variations in chromosomes, genitals, and internal reproductive organs. Intersex variations are naturally occurring and not a disorder, disease, or condition [50].

Taken together, these perspectives highlight that gender is not a fixed property that can be simply determined from a body or a voice, but an interpretation shaped by social norms and repeated categorisation practices. In HRI, this matters because robot gender is likewise produced through users' inferences from design cues and interaction context. The next subsection, therefore, presents design strategies aimed at reducing gender attribution in robots, before turning to voice and discussing how it can be used to avoid gendering and stereotyping.

2.1.3 Reducing Gender Attribution in Robots

Given the complexities of robot gendering and the persistence of the masculine-by-default effect, researchers have begun exploring design strategies aimed at reducing gendering. One such approach is the use of gender-neutral designs.

The term "gender-neutral" generally refers to the absence of explicit gender cues, with the goal of preventing users from assigning gender and, in turn, reducing the negative effects of gendering [53, 54]. Within HRI, this approach is often motivated by the assumption that removing gender indicators from an agent's design may limit gender-based expectations and biases. Empirical work provides some support for this idea. For example, Warta [55] found that participants preferred a human-like, gender-neutral robot over explicitly gendered alternatives. Similarly, Otterbacher et al. [56] showed that robots lacking overt gender cues did not elicit uncanny valley responses, while Strait et al. [57] observed that gender-neutral robots received fewer negative comments than gendered ones. Together, these studies suggest that reducing explicit gender cues can have positive effects on user perception.

However, the notion of gender neutrality is not without limitations. Sutton [22] argues that artificial agents cannot be truly gender-neutral, as humans tend to assign gender whenever human-like features are present. Supporting this claim, research shows that visual cues such as hair length can trigger gendered attributions [4], vocal characteristics strongly shape perceived gender [40], and even the occupational context associated with an agent influences how it is gendered [41]. These limitations suggest that removing gender cues entirely may be neither feasible nor sufficient, motivating alternative approaches to reducing gender attribution.

In this context, voice emerges as a particularly important design dimension. A recent review by Seaborn et al. [58] highlights voice as a key factor influencing gender perception in robots and artificial agents. Simply hearing a voice is often sufficient for users to assign gender, shaping expectations and interpretations of the robot [26, 58, 59, 60]. This central role of voice is also reflected in HRI research practice. In their review, Perugia and Lisy [45] report that in the majority of the studies they examined, 78% of researchers relied on voice as the primary cue to manipulate a robot's gender. Precisely because voice plays such a strong role in gender attribution, it has become a key focus for design interventions aimed at reducing gendering. As a result, there is a growing interest in the development of non-gendered voices, both in research and industry, enabled by advances in text-to-speech technologies (TTS) [9, 61, 62]. These voices represent a promising direction for more inclusive and less stereotyped/gendered interactions with technology [63]. In the next section, I first discuss the role of voice in HRI, why it is such a powerful design dimension, and then move to gender-ambiguous voices and their potential.

2.2 The Role of Voice in HRI

Voice is one of the primary channels through which humans communicate and understand each other, and it therefore plays a central role in shaping social interaction. The human voice does more than simply convey spoken content: it also carries rich information about who is talking, such as gender, age, emotional state, and origin. People then draw on these cues to quickly form impressions of a person, including whether they like and trust them [64]. This is possible because voice is composed of multiple attributes, such as volume, timbre, pitch, rhythm, and articulation, all of which contribute to how vocal information is perceived and interpreted. Among these attributes, pitch, defined as the perceptual correlate of the voice's fundamental frequency (F0), is a very easily quantifiable characteristic that influences many aspects of person perception, including attraction, trustworthiness, and charisma. For example, prior research shows that women tend to perceive higher-pitched male voices as less attractive than lower-pitched ones [65], while both women [66] and men [67] generally prefer higher-pitched female voices. Related work also shows that other vocal cues shape these broader judgements: accent and prosody can influence trusting behaviour [68], and acoustic patterns such as melody, tempo, and fluency contribute to perceived charisma [69].

These findings underline the central role of voice in social perception. As a result, voice has also been extensively explored in human-technology interaction. Within the field of HRI, a growing body of work demonstrates that voice strongly affects how users perceive and interact with robots. Early work by Niculescu et al. [70] showed that vocal pitch alone can shape robot evaluations: a high-pitched voice led people to perceive a robot as more appealing, emotional, and engaging, whereas a low-pitched voice robot was associated with strength and rationality, but lower social appeal. More recent work has shifted attention to the increasing realism of synthetic voices and their implications for interaction. For example, Becker et al. [71] compared natural-sounding and mechanical robot voices and found that users expect robots to sound human-like and perceive robots with natural voices as safer. While natural voices did not directly increase reported trust, they influenced user behaviour by increasing compliance, even when the robot provided incorrect advice. Together, these studies show that vocal characteristics shape not only how robots are perceived, but also how people respond to and rely on them during interaction.

Beyond individual vocal attributes, prior work has shown that coherence between a robot's voice and its appearance is essential, and the quality of interaction can depend on it. Mara et al. [72], for example, demonstrated that human-like voices prompt people to imagine more human-like robot features, such as a nose or hair, whereas less human-like voices lead to more mechanical designs, such as robots with wheels. McGinn and Torre [60] investigated the mental images participants form when hearing robots speak and found that even when speech is incomprehensible, participants still associate voices with specific robot forms. Their findings suggest that both voice and appearance shape robot perception, and that assigning an incongruent voice may introduce a confounding factor in HRI studies. Extending this line of work, Torre et al. [73] explored how participants match voices to robot images across different contexts. They found that the same vocal characteristic could lead to different robot choices depending on the situation, indicating that people hold expectations about how a robot should sound even before hearing it.

Taken together, this literature shows that voice is not a secondary design feature but a primary driver of robot perception, shaping users' mental models. Because voice influences impressions, expectations, and interaction quality from the earliest stages, it requires careful design consideration. In the next section, I introduce voices beyond the binary, present the state of the art, and discuss their potential to reduce gendering and support inclusion.

2.3 Voices Beyond the Binary

Voices beyond the binary offer a concrete design space for testing whether gender attribution and stereotyping can be reduced. These voices are variously described as gender-neutral, genderless, gender-free, ambiguous, nonbinary, or gender-expansive [22, 24, 74, 75]. While the terminology varies, these efforts share a common goal: to move beyond binary gender categories while maintaining clarity and naturalness. Offering alternatives beyond male and female voices is relevant not only for reducing stereotyping but also for inclusivity. Gender-diverse and gender-nonconforming individuals may feel better represented when technology does not force alignment with binary gender categories [13, 18].

A growing body of research in HRI and HCI has begun to empirically investigate voices that move beyond binary gender categories. One of the first studies in this area was conducted by Torre et al. [14], who examined whether a gender-ambiguous voice could interfere with gendering when other robot characteristics were strongly gendered. In a video-based study, participants evaluated robots whose appearance and occupational roles were stereotypically masculine or feminine, either with or without a gender-ambiguous synthetic voice. Their results showed that adding a gender-ambiguous voice reduced participants' tendency to gender the robot, even when its body and role were gendered. Importantly, the ambiguous voice did not increase uncanniness or reduce trust, suggesting that ambiguity can act as a subtle counterbalance to other gendered cues rather than as a disruptive element.

Complementing this work, Längle et al. [13] investigated user perceptions of male, female, and nonbinary voices in the context of digital voice assistants. In a large-scale questionnaire study, participants listened to different voices and evaluated perceived gender, personality traits, and likeability. While binary voices elicited clear gender-stereotypical associations, the nonbinary voice did not trigger stereotypical masculine or feminine trait attributions. Although the nonbinary voice was often perceived as leaning slightly masculine, participants did not associate it with gendered personality traits, suggesting that gender perception and stereotyping did not align.

Beyond perception studies, recent work has also explored how nonbinary and gender-expansive voices can be designed and evaluated in collaboration with gender-diverse communities. Danielescu et al. [9] presented a method for creating a nonbinary text-to-speech voice developed with input from nonbinary and transgender participants. Their large-scale evaluation showed that nonbinary users were significantly more likely to prefer a nonbinary voice than cisgender users, highlighting the importance of representation and choice rather than assuming a single neutral solution. Similarly, Székely and Hope [63] proposed a palette-based approach to gender-expansive synthetic voices, trained on recordings from gender-expansive speakers and evaluated with nonbinary users of speech-generating devices. Their findings emphasise that gender-expansive voices are not a single category but a space of variation, shaped by multiple acoustic and expressive dimensions.

Taken together, this body of work shows that investigating voices beyond the binary is both theoretically and practically important. It addresses the limitations of binary gender frameworks in technology, highlights the social impact of voice design choices, and opens new directions for inclusive and reflective HRI research. In particular, voices beyond the binary have the potential to disrupt habitual processes of gender attribution by resisting clear placement within masculine or feminine categories. The mentioned challenges motivate the research presented in this thesis. In particular, there is still limited understanding of how people perceive voices beyond the binary (RQ1), whether such voices meaningfully increase feelings of inclusion (RQ2), and to what extent they can reduce gender attribution in human–robot interaction (RQ3).

In the next chapter, I summarise the four papers included in this thesis and describe how they build on one another to form the empirical and conceptual foundation of this work.

Chapter 3

Summary of the Included Papers

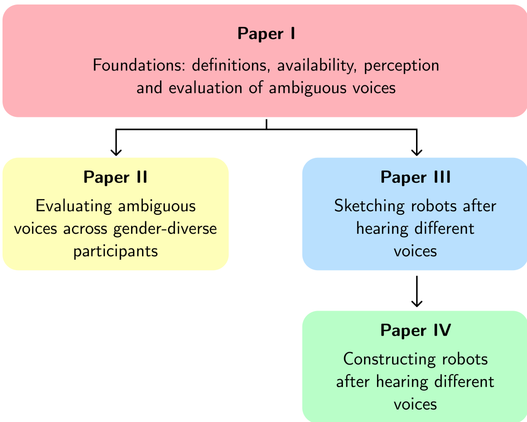

Figure 3.1: Overview of the included papers and how they build on one another.

This chapter summarises the four papers included in this thesis. Figure 3.1 outlines the thesis trajectory: Paper I establishes the conceptual baseline, Paper II tests perception across listener groups, and Papers III–IV examine how voice cues translate into robot gendering in both sketches and physical prototypes.

Paper I marked the starting point of my PhD. I conducted a systematic review on gender-ambiguous voices to clarify the state of the art and identify the research gaps that guided the rest of my work. Paper II examines how ambiguous voices are perceived by participants of different gender identities, with equal numbers of women, men, and nonbinary people. Paper III examines whether ambiguous voices can reduce robot gendering, using a sketching-based method to see how participants draw robots after listening to different voices (masculine, feminine, ambiguous). Paper IV builds on Paper III through a follow-up study. Using 3D robot body parts created from earlier sketches, I asked participants to assemble a robot prototype after they listened to either a masculine, feminine, or ambiguous voice to further examine robot gendering. In the following subsections, I provide a detailed summary of each paper.

3.1 Paper I - Breaking the Binary: A Systematic Review of Gender Ambiguous Voices in Human-Computer-Interaction

As mentioned above, [19] allowed me to explore the area of gender-ambiguous voices and to establish the base for my PhD journey. To do so, I reviewed 36 articles in the field of HCI, investigating the definitions and availability of gender-ambiguous voices, their creation methods, user perception, and evaluation techniques. This research was guided by three main research questions:

- **RQ1:** What defines an ambiguous voice in HCI literature?
- **RQ2:** What are the currently available gender-ambiguous voices?
  - **RQ2-a:** What is the availability of gender-ambiguous voices?
  - **RQ2-b:** What characteristics do gender-ambiguous voices have?
  - **RQ2-c:** What methods were utilised to create these gender-ambiguous voices?
- **RQ3:** What is the perception of gender-ambiguous voices?
  - **RQ3-a:** What methods were used to evaluate the perception of gender-ambiguous voices?

For **RQ1**, I found that most papers used terms such as "gender-neutral" or "gender-ambiguous", yet only five papers provided definitions, leading to inconsistent terminology. In response, I clarified key terminology and proposed distinctions between terms such as gender-neutral, gender-ambiguous, and nonbinary.

For **RQ2** and **RQ2-a**, I distinguish between identifying which beyond-the-binary voices are reported in the literature and whether those voices are actually open-source. Voice availability was often unclear: most papers did not specify whether their voices were open source, and only three explicitly stated that they were. To address this gap, I created an open-source repository of gender-ambiguous voices. For **RQ2-b**, I contacted the authors of the 36 papers to obtain audio samples, collecting four voices that I analysed acoustically. The voices varied widely in pitch, resonance, and voice quality, indicating that ambiguity cannot be achieved through pitch adjustment alone. I used these findings to propose initial recommendations and a preliminary framework for creating and using ambiguous voices. Finally, for **RQ2-c**, the reviewed papers used diverse approaches to produce ambiguity, including commercial

TTS voices, customised synthetic voices, modulated human voices, and, in a few cases, newly developed TTS systems.

For **RQ3**, the review showed mixed perceptions. Some studies reported negative reactions, while others found more positive responses, especially among nonbinary participants and sometimes women. Overall, the findings showed no clear consensus. For **RQ3-a**, evaluation methods varied widely. Studies used different rating scales, or qualitative analysis, making comparisons difficult and showing the need for more standardised approaches. To address this, I proposed an initial 5-point Likert scale to evaluate gender-ambiguous voices.

This paper contributes to my licentiate by helping me understand the state of the art, identify key conceptual, methodological, and practical gaps for the following papers, and establish the research context for my later contributions. Beyond this thesis, it contributes to the HCI and HRI community by clarifying and defining the terminology used for voices beyond the gender binary (including gender-ambiguous, gender-neutral, genderless, nonbinary, etc.), presenting an exploratory acoustic analysis of a small set of gender-ambiguous voices, proposing concrete guidance and a dedicated evaluation scale for their assessment, releasing a public repository of available gender-ambiguous voice samples, and offering a structured framework to support researchers in creating, selecting, and using voices beyond the binary in empirical studies.

3.2 Paper II - Hearing Ambiguity: Exploring Beyond-Gender Impressions of Artificial Ambiguous Voices

Taking inspiration from the works reviewed in Paper I, particularly Danielescu et al. [9], which emphasise the importance of involving gender-diverse users, I identified a clear gap: studies on artificial non-gendered voices rarely include participants from underrepresented groups, especially queer and nonbinary people. For this reason, in this study [76], I examined how six ambiguous voices are perceived by women, men, and nonbinary participants (74 per group). I also included one masculine and one feminine voice as attention checks. Perception was measured across the five VOXI dimensions [77]: trustworthiness, appeal, comfort, anthropomorphism, and aversion. Specifically, in this online experiment, I investigated:

> **RQ1:** How do participants evaluate ambiguous voices across different voice perception dimensions?
>
> **RQ2:** Does priming participants on the gender ambiguity of artificial voices affect their perception of such voices?
>
> **RQ3:** Does participant gender influence their perception of artificial gender-ambiguous voices?

For **RQ1**, the results showed that participants evaluated the six ambiguous voices differently across the five VOXI perception dimensions. Although some ambiguous voices were consistently rated more positively, others (e.g., Sam from [9]) were perceived as less trustworthy or more aversive. The two gendered voices (one male and one female) were generally rated more positively than most ambiguous voices, indicating that listeners often still prefer gendered voices. Together, these findings show that ambiguity alone is not sufficient for creating universally appealing voices and that acoustic and sociolinguistic factors, such as accent, strongly shape impressions.

For **RQ2**, priming participants about the ambiguity of the voices did not significantly shift overall perceptions. However, small interaction effects showed that priming made nonbinary participants slightly more critical, lowering ratings in dimensions such as trustworthiness and comfort. This suggests that priming may increase awareness, particularly for listeners for whom gender ambiguity might be personally relevant.

For **RQ3**, participant gender significantly influenced perceptions of the voices. Men and women generally provided similar evaluations across all dimensions. In contrast, nonbinary participants rated the ambiguous voices more negatively overall, giving lower scores for trustworthiness, appeal, comfort, and anthropomorphism, and higher scores for aversion. These findings indicate that currently available ambiguous voices do not fully resonate with nonbinary listeners, highlighting a mismatch between design intentions and user experience.

This paper contributes to my licentiate by helping me understand how perceptions vary across gender-diverse listeners and by identifying factors that may limit the inclusivity of ambiguous voices, highlighting the importance of involving queer and nonbinary communities in the development of inclusive voice technologies. Beyond this thesis, it contributes to the HCI and HRI community by providing one of the first large-scale, comparative evaluations of multiple gender-ambiguous voices across several perceptual dimensions, with balanced participation from women, men, and nonbinary listeners. By moving beyond gender classification and examining dimensions such as trustworthiness, comfort, appeal, anthropomorphism, and aversion, this work broadens how gender-ambiguous voices are evaluated and perceived, and provides empirical grounding for the design of more inclusive voice technologies.

3.3 Paper III - Sketching Robots: The Impact of Gender Ambiguous Voices on Mental Images of Robots

Paper III [78] extends this thesis from how gender-ambiguous voices are perceived to how they shape people's mental models of robots. It was also my first empirical study on robot gendering. In this work, I examined whether a gender-ambiguous voice affects how people imagine a robot across different occupational contexts. The study used a mixed-method approach combining sketching and qualitative analysis. Participants listened to one of several voice-scenario combinations (masculine, feminine, or ambiguous) paired with a stereotypically masculine, feminine, or neutral occupational scenario (e.g., security guard, secretary, tour guide). After listening, they sketched the robot they imagined and described the factors that influenced their sketch. The work was guided by the following research question:

> **RQ:** Does a gender-ambiguous voice influence how people gender a robot in different occupational contexts?

The results showed that a gender-ambiguous voice influenced gender associations. Across conditions, participants were less likely to assign a gender to their sketched robots when exposed to an ambiguous voice compared to a masculine or feminine voice. These findings align with and extend existing work on the influence of ambiguous voices in HRI [14], providing early evidence that such voices can reduce robot gendering even in highly gendered occupational settings. At the same time, a masculine-by-default tendency was visible: in the neutral scenario (tour guide) with the ambiguous voice, several participants still sketched or described their robot as male, reflecting well-documented biases in HRI that position male as the assumed default for robots [48].

This paper contributes to my licentiate by building directly on the gaps identified in Paper I. While Paper I examined how gender-ambiguous voices are perceived in existing studies, this study extends that work by exploring perception through the mental images people form when imagining a robot. It also marks an important step in my research journey, as it represents my first empirical study on robot gendering and my first use of sketching as a research method, allowing me to move from summarising prior work to designing and running novel empirical investigations. Beyond this thesis, it contributes to HRI research by introducing sketching as a method for examining how vocal cues shape gendered mental models of robots, and by providing a curated dataset of 180 robot sketches that surfaces implicit gender assumptions difficult to capture through traditional rating or matching approaches.

3.4 Paper IV - From Voice to Form: How Gender-Ambiguous Voices Shape Physical Robot Design

Building on Paper III, [79] examines how ambiguous, feminine, and masculine voices influence the physical design choices people make when constructing a robot prototype. In this mixed-method, in-person study, participants listened to a robot introduce itself using an ambiguous, feminine, or masculine voice in a neutral occupation (tour guide), built a physical prototype using a toolkit of modular body parts (heads, torsos, limbs), took part in an interview about their design process, and concluded by evaluating both the voice and the robot they had created. The following research questions guided this work:

RQ1: Does a voice (ambiguous, masculine, or feminine) influence participants' design choices during robot construction/customisation?

RQ1-a: Does the voice type influence whether participants build robots with more human-like or machine-like characteristics?

RQ1-b: Does the voice type influence whether participants build robots with gendered characteristics?

For **RQ1**, I found that voice did influence participants' design choices, even when this influence was not always reflected in their explicit questionnaire ratings. Indeed, participants' prototypes and interview explanations showed that they adapted their designs based on how they perceived the robot's voice, suggesting that the voice acted as an early design cue during construction.

For **RQ1-a**, although participants' explicit ratings did not show differences in anthropomorphism, the prototypes and interview data did. Robots built after hearing the ambiguous voice contained fewer human-like pieces and more hybrid or machine-like elements. In contrast, the masculine voice produced more anthropomorphic robots with human-like components. The feminine voice did not lead to a clear pattern in anthropomorphism.

For **RQ1-b**, the masculine voice led participants to use more masculine pieces, while the feminine voice encouraged the use of feminine ones. The ambiguous voice led to fewer gendered pieces, although a masculine-by-default tendency remained, with some participants still selecting masculine parts. This mirrors the findings from Paper III and suggests that ambiguous voices can reduce, but not entirely remove, gender attribution during physical design.

This paper contributes to my licentiate by extending the findings of Paper III from sketched robot form to physical robot construction. By examining how gender-ambiguous, feminine, and masculine voices influence participants' concrete design choices during robot prototyping, it deepens my understanding of how voice shapes embodiment and gender attribution beyond mental imagery. Together with Paper III, these findings highlight the importance of beginning the robot design process with voice rather than treating it as an add-on and suggest that using gender-ambiguous voices may help reduce gendering in robot design. Beyond this thesis, it contributes to the HRI community by demonstrating how voice influences physical robot embodiment, robot gendering, and by providing the RoMix toolkit as an open, reusable resource for studying gender, anthropomorphism, and embodiment through hands-on robot prototyping.

Together, these four papers helped me answer my research questions and shaped my understanding of the future work I want to pursue. In the next chapter, I explain how the findings from these papers contribute to answering my PhD research questions, outline my contributions, and show how these studies will inform my future research directions.

Chapter 4

Discussion and Future Directions

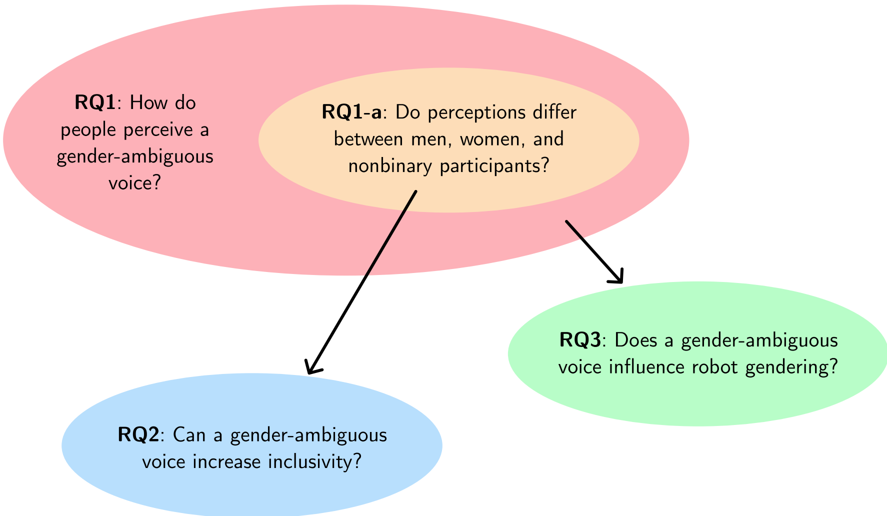

Figure 4.1: The diagram shows how the research questions guiding this thesis are interconnected. RQ1 was the starting point of my PhD, with RQ1-a as a more detailed part of it. Arrows show how RQ1-a leads to RQ2 and how RQ1 connects to RQ3.

So far, this thesis has laid out the background, motivations, and research trajectory behind my work. It is now time to return to the questions that guided these two years of my PhD, both the ones I started with and the ones that emerged along the way. Figure 4.1 illustrates how these research questions relate to one another.

I begin with the first question I established at the very start: how people perceive gender-ambiguous voices (RQ1). I then move to a question that surfaced during the first year: whether people of various gender identities perceive ambiguous voices differently (RQ1a). Next, I address a question that arose at the beginning of my second year: whether gender-ambiguous voices can increase inclusivity (RQ2). Finally, I end with the question that stayed with me throughout the entire project: whether a gender-ambiguous voice influences how people gender a robot (RQ3).

The answers to these questions, together with all the data I collected across these two years, shape the contributions of this thesis and guide my future directions, which I outline in Section 4.6 after addressing each research question.

4.1 RQ1: How Do People Perceive a Gender-Ambiguous Voice?

To answer this research question, it is important to present the insights I gained from three of my studies: Paper I (the literature review [19]), Paper II (the voice perception study [76]), and Paper III (the sketching study [78]).

In Paper I (the literature review [19]), I examined how previous work described and evaluated gender-ambiguous voices. Across the 36 papers, researchers used many different labels for these voices: "gender-neutral", "gender-ambiguous", "nonbinary", "genderless", and more, often without defining them consistently. This lack of agreement around terminology reflects a broader issue in the field: perceptions of these voices vary widely. Several studies reported negative reactions, with participants describing ambiguous voices as annoying, artificial, or less desirable than gendered options [9, 23, 80]. At the same time, some groups responded more positively. Nonbinary listeners often evaluated these voices as more trustworthy or more representative, especially when the voice carried slightly feminine cues such as higher intonation [9]. A similar feminine pull appeared in Torre et al. [14], where listeners tended to perceive an ambiguous voice as leaning feminine even when ratings sat near the midpoint.

Perceptions also varied across listener demographics. Tolmeijer et al. [7] found that women rated the ambiguous voice as more friendly, polite, and trustworthy than men did. Other studies showed that many listeners struggled to assign a gender at all, and a notable number even preferred the ambiguous voice to a gendered one [81]. Voice type mattered as well: natural, ambiguous voices were evaluated more favourably than synthetic ones [82], suggesting that naturalness reduces some of the uncertainty associated with ambiguity.

Overall, my review shows that there is no single perceptual pattern for gender-ambiguous voices. Impressions shift with acoustic cues, listener expectations, demographic factors, and the naturalness of the voice. This variability sets the foundation for Papers II and III, where I explore these perceptions in more controlled and focused ways.

Paper II (the voice perception study [76]) provides more details to answer to RQ1 by examining how six ambiguous voices (collected through the systematic review), alongside masculine and feminine voices used as attention checks, were evaluated across five VOXI perception dimensions [77]: trustworthiness, comfort, appeal, anthropomorphism, and aversion. Across the full sample, the ambiguous voices were not perceived uniformly. Some were evaluated more positively, while others elicited higher aversion or lower trust. Importantly, the masculine and feminine voices consistently scored higher than most ambiguous voices on dimensions such as trust, comfort, and appeal, indicating that listeners tended to prefer clearly gendered voices. This suggests that gendered voices remain a perceptual "default" even when ambiguity is available as an alternative. It also reflects a broader pattern identified in my systematic review: ambiguity can disrupt gendering, but it does not necessarily produce a better listening experience. The perception of a voice as ambiguous does not guarantee that it is perceived as natural, pleasant, or trustworthy.

Behind the dimensions explored in Paper II, it was essential for me to understand whether people actually perceive ambiguous voices as such. For this reason, in Paper III (the sketching study [78]), among other things, I ran a validation study focused directly on ambiguity perception. Here, 180 participants rated the same six ambiguous voices as before on whether they sounded feminine, masculine, or ambiguous. When evaluating ambiguity, I gave them the following definition: *"by Ambiguous", we mean that the voice sounds neutral or androgynous"*. Three of the six voices were consistently perceived as ambiguous (see Figure 4.2), but only one (Sam, from Danielescu et al. [9]) emerged as the most balanced (also when evaluating perceived masculinity/femininity of that voice), and it was then the voice used in the rest of the experiment. This step was crucial because it showed that ambiguity is not a property of the sound itself but an outcome of how people hear and classify it.

Overall, the findings from Papers I, II and III answer RQ1: perception of ambiguous voices is multidimensional, uneven, and shaped by a combination of acoustic, sociolinguistic, and cultural factors. As shown, ambiguity might influence gendering, but it does not necessarily improve user experience. Given these results, it became clear to me that I needed to delve deeper into how the perception of ambiguous voices changes across different groups of listeners. This, then, led to the next research question, which focuses specifically on the role of participant gender and their perception of ambiguous voices.

4.1.1 RQ1a: Do Perceptions Differ Between Men, Women, and Nonbinary Participants?

To answer RQ1a, we need to look into Paper II (the voice perception study [76]) where as participants I had a balanced sample of men, women, and nonbinary people (74 per group). The results showed clear differences among groups.

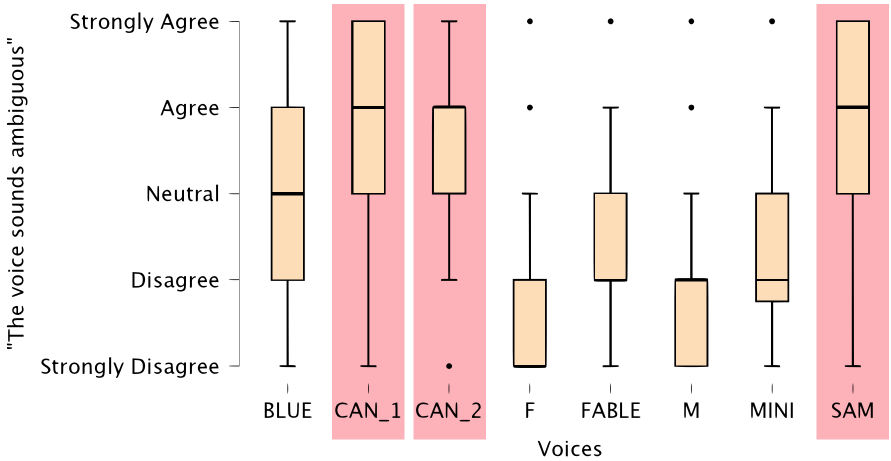

Figure 4.2: Distribution of ratings given in the ambiguous statement for all eight voices (F=female voice, M=male voice).

Men and women evaluated the ambiguous voices similarly across all perception dimensions, generally providing moderate to positive evaluations. Nonbinary participants, however, consistently evaluated the voices more negatively, reporting lower trustworthiness, comfort, appeal, and anthropomorphism, and higher aversion.

To help understand these results, it is important to clarify the study design. Half of the participants were primed, meaning they were explicitly told before the experiment that the voices they would hear were gender-ambiguous and were given a short definition of what this meant. The remaining participants were not primed and received no information about ambiguity; they were simply instructed to listen to the voices. Across both the primed and non-primed conditions, nonbinary participants evaluated the voices more critically than men and women. Priming further intensified this effect: when ambiguity was explicitly mentioned, nonbinary listeners expressed even lower trust and comfort, and higher aversion.

Taken together, the results from Paper II provide a clear answer to RQ1a: perceptions of gender-ambiguous voices differ significantly across gender groups. Before running this study, I had almost assumed that, because these voices might reduce gendering and move away from the usual masculine and feminine options, they could also serve as an inclusive choice for people outside the binary. Based solely on Paper II, however, this does not seem to be the case. Ambiguous voices lowered trust and comfort and increased aversion among nonbinary participants. This came out as a surprise since Danielescu et al. [9] found the opposite pattern, with nonbinary listeners responding more positively to ambiguous voices. This contrast led me to the next research question.

4.2 RQ2: Can a Gender-Ambiguous Voice Increase Inclusivity?

This research question is still a work in progress and represents an ongoing direction for my future work. Part of its answer appears in Paper II (the voice perception study [76]), where one finding is that ambiguous voices do not create the same experience across gender groups, with nonbinary people being more negative about them. Ambiguity, then, as it is currently designed, does not automatically translate into inclusion, recognition, belonging, or comfort for the population these voices are often intended to support. One possible explanation for this pattern is that nonbinary listeners may feel more directly addressed or affected by gender-ambiguous voices, which are often positioned as an inclusive alternative to binary gendered voices. In this context, the label "ambiguous" itself may have been experienced as problematic, as it might have been interpreted as implying that nonbinary identities are vague or indeterminate, potentially leading to feelings of rejection rather than recognition. As a result, nonbinary listeners may approach these voices more critically, holding them to higher expectations in terms of authenticity, comfort, and representation.

Despite this, I believe ambiguous voices can still support inclusivity, but in a different way. They introduce uncertainty in the listener. Indeed, every time I play an ambiguous voice in a room, and I ask who perceived the voice as masculine and who as feminine, the room is usually divided. As an example, in Figure 4.3, I show a result from a short activity with some of my students. I asked them how they perceived a voice without mentioning anything about my research on ambiguous voices, and without including "ambiguous" as an option in the response scale. As we can see, the students tended to position the same voice as masculine or feminine or something else. In this sense, then, I believe ambiguity can increase inclusivity by preventing normative assumptions. Inclusivity here is not about resonance with identity, but about reducing biases.

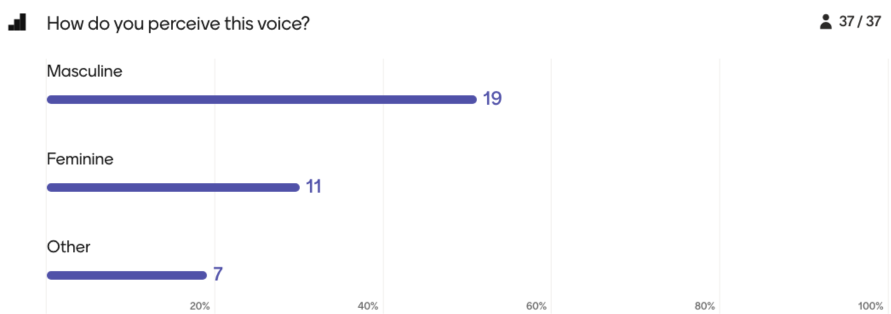

Figure 4.3: 37 students' ratings to the question "How do you perceive this voice?" after I played an ambiguous voice in one of my classes.

Of course, this form of inclusivity is completely different from identity-based inclusion. In my opinion, identity-based inclusion means creating something **for** a specific community and creating it **with** that community [83]. In a currently under-review paper, I argue that the term “ambiguous” should not be used to represent queer and nonbinary people’s voices in HCI and HRI research. While ambiguous voices play an important role in reducing categorisation and gendering, they are not sufficient when the goal is the inclusion of gender non-conforming people. In such cases, labels such as “queer voice” or “nonbinary voice” should be considered, as they more accurately represent the communities they aim to include. In the under-review paper, I also present a survey with nonbinary participants, who were asked to describe synthetic voices created from nonbinary speakers (see [84] for details on how these voices were developed). To describe the voices, they used terms such as “nonbinary” or “any gender” rather than “ambiguous”. Figure 4.4 shows a word cloud illustrating the distribution of voice labels used by participants. This work helped me understand that I, and other researchers, should avoid describing ambiguous voices as inclusive of gender non-conforming people. The empirical findings from Paper II also support this view. When nonbinary participants were told that the voice was ambiguous and were given a definition of it, they reported lower trust and comfort and higher aversion. This result aligns with the findings of the under-review paper, suggesting that participants did not feel represented by the term ambiguous and therefore distanced themselves from it.

Figure 4.4: Word cloud showing the distribution of labels used by nonbinary participants to describe synthetic voices created from nonbinary speakers.

These two pieces of work helped answer RQ2: ambiguous voices can increase inclusivity, but not by creating inclusive experiences for queer and nonbinary people. Ambiguous voice contribution lies in reducing gendering, as I discuss in the next section, and in making technology less binary by default. They do not create inclusivity at the level of identity because they do not offer recognition or representation. Inclusivity in terms of user experience requires naming and designing voices in ways that reflect the identities of queer and gender diverse users, something that ambiguous voices were never intended to do.

What I want to highlight here is the need for community-informed design. Inclusivity in terms of user experience requires collaboration with nonbinary people, both to shape future voice palettes and to determine which labels are appropriate for different design goals. Ambiguity is one tool, but it is not sufficient on its own. Designing voices that feel genuinely inclusive will require moving beyond ambiguity. It also means using labels and categories grounded in community experience, such as "nonbinary" or "any gender". At the same time, ambiguity can still be useful in contexts where the goal is to reduce gendering rather than to represent specific users. The next research question investigates this more directly by examining the effect of ambiguous voices on robot gendering.

4.3 RQ3: Does a Gender-Ambiguous Voice Influence Robot Gendering?

Paper III and IV help answer this research question. Both studies investigated how an ambiguous voice can influence robot gendering, the first through an online sketching study and the second through an in-person prototyping study.

In Paper III (the sketching study [78]), 180 participants each listened to a single voice–scenario combination. Three scenarios were used (masculine, feminine, and neutral) based on a prior validation study. Each scenario was paired with either a gender-ambiguous, masculine, or feminine voice. Participants listened to one of these short narrated scenarios and then had 10 to 15 minutes to sketch the robot they imagined having that voice. This between-subjects design allowed me to capture first impressions by ensuring that each participant reacted to only one voice and one scenario at a time. After producing the sketch, participants answered free-text questions about different aspects of their drawing, including whether they sketched the robot with a specific gender in mind. The sketches were then analysed for gender cues such as body shape, hair, facial features, and clothing.

The results showed a strong voice-gender connection in the gendered-voice conditions: feminine voices led to exclusively feminine sketches, and masculine voices led to exclusively masculine ones. In contrast, participants who heard the ambiguous voice produced sketches that were mostly coded as neutral or that mixed masculine and feminine elements rather than aligning clearly with one category. This showed that the ambiguous voice has the power to weaken binary expectations at the level of imagined form. This can also be seen in the sketches shown in Figure 4.5.

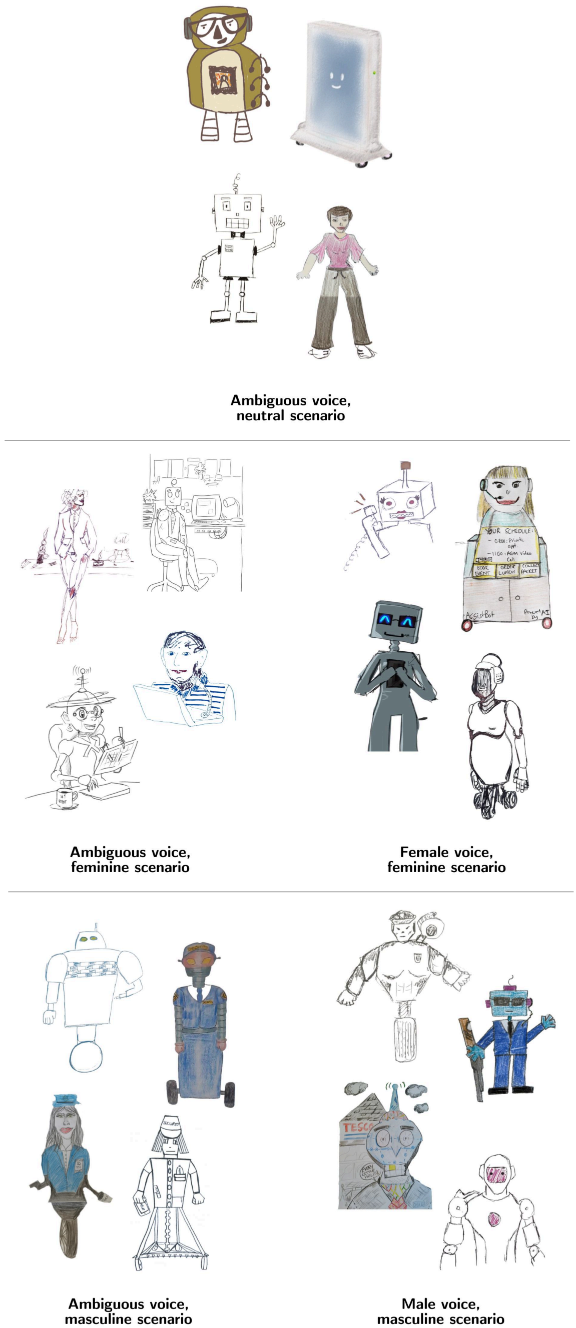

Figure 4.5: Distribution of the appearance of some sketched robots across the five voice–scenario conditions.

Even when an ambiguous voice was paired with a masculine or feminine scenario, participants often mixed characteristics of both genders. For example, in the first sketch in the top left of the "ambiguous voice, feminine scenario" condition (see Figure 4.5), the drawing has a feminine appearance but includes a beard, a more masculine jawline, and a suit typically associated with masculinity.

In Paper IV, the prototyping study [79], which builds on Paper III, I further examine how ambiguity affects embodiment. This study moves from 2D sketches to 3D physical prototype construction. Participants assembled a robot using a set of modular parts that varied in both anthropomorphism (human-like, human-machine-like, machine-like) and gender (male, female, neutral). The body parts were modelled on the sketches collected in Paper III. Figure 4.6 illustrates this process by showing an example of the masculine head set, in which participants' sketches were transformed into 3D-printable parts. Figure 4.7 shows the resulting RoMix toolkit used in the study.

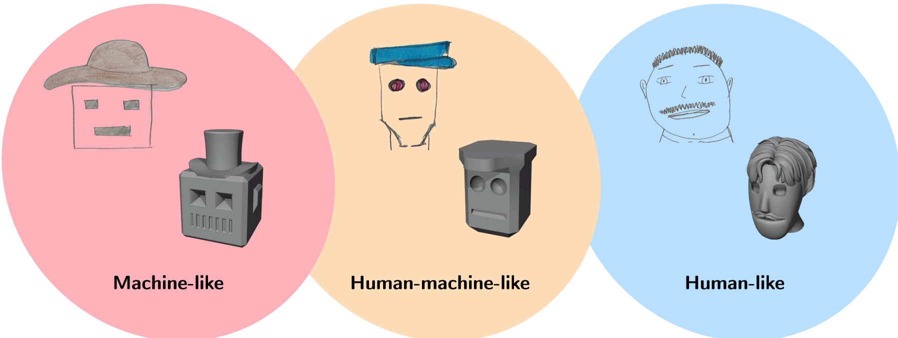

Figure 4.6: Examples of masculine head stimuli. Top: sketches from Paper III [78]. Bottom: corresponding 3D models.

Figure 4.7: Experiment setup showing a participant assembling a robot using the provided RoMix toolkit. The inset highlights the box containing interchangeable heads, torsos, arms, hands, legs and feet used in the study.

Using physical robot construction as an approach made it possible to observe more concrete and constrained choices and to assess whether the patterns found in the sketches remained when participants had to commit to actual physical components. The results show that the ambiguous voice did not eliminate gendering entirely, but rather changed how gendering appeared. When listening to the ambiguous voice, participants constructed fewer human-like robots and selected more hybrid or machine-like parts, often combining features that resisted clear categorisation. However, even when participants labelled their robot as neutral, the prototypes still contained subtle gender cues. Notably, those exposed to the ambiguous voice tended to select more masculine torsos than participants in the feminine voice condition, suggesting that people move away from feminine cues more easily than from masculine ones.

Figure 4.8 shows examples of the robot prototypes built in each voice condition. Prototypes created after hearing the feminine voice exhibit clearly feminine characteristics, while those built in the masculine voice condition display masculine features. In contrast, robots assembled in the ambiguous voice condition include a mix of different parts, although most selected torsos were masculine.

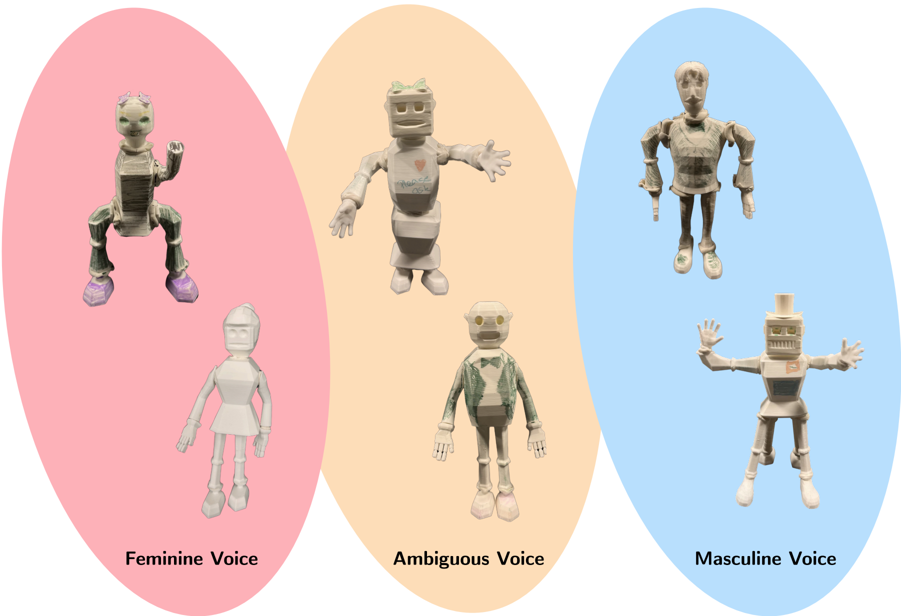

Figure 4.8: Six prototypes built by participants, divided by voice condition.

Taken together, these findings answer RQ3: a gender-ambiguous voice does influence robot gendering, but it does so by softening gender cues rather than eliminating them. Ambiguity destabilises the binary, yet masculine defaults persist, showing both the potential and the limits of this approach within broader cultural patterns that shape how people imagine and construct robots.

Extending this perspective across the studies reviewed and conducted in this thesis, a broader pattern emerges. Gender-ambiguous voices can reduce gender attribution, opening up design space for more inclusive, flexible, and context-aware interactions with robots and conversational agents. At the same time, these approaches come with important challenges. Gender perception depends on listeners' cultural background, gender identity, age, and experience, and ambiguous voices are often still interpreted through binary expectations [22, 53]. Moreover, voice is only one of several cues through which gender is assigned: language use, personality traits, visual design, and interaction context can override vocal ambiguity. Gender-ambiguous voices should therefore not be treated as a stand-alone solution, but as one element within a broader, gender-sensitive interaction design.

4.4 Summary of My Contributions

Across these four papers, my work advances research on gender-ambiguous voices in HCI and HRI in ways that differ substantially from prior literature. Indeed, many studies use ambiguous (neutral, genderless, or similarly labelled) voices without systematically examining whether these voices are perceived as such, how they are experienced across different user groups, how they affect inclusion, or how they shape robot design. My contributions address these gaps across four levels: conceptual, methodological, empirical, and practical.

1. **Defining ambiguity** ***(Conceptual)***: Through the systematic review (Paper I), I show that the HCI field uses many different and inconsistent terms for non-gendered voices, and that only a few studies define these terms clearly. I propose "gender-ambiguous" as a more accurate label when the aim is to reduce gendering, and show that ambiguity is not a fixed property of the voice but something listeners construct based on their expectations, backgrounds, and cultural context. This means that ambiguity must be tested and validated, not assumed. Earlier work rarely questioned the terminology used, and this thesis offers the first clear critique of this issue.

2. **Evaluating ambiguous voices across genders** ***(Empirical)***: Paper II contributes the first study that evaluates six ambiguous voices using a validated multi-dimensional instrument (VOXI) [77], with balanced samples of men, women, and nonbinary participants. My findings reveal that nonbinary participants evaluate ambiguous voices more negatively than men and women. This challenges the common assumption that ambiguous voices "naturally" support inclusivity. It shifts the conversation from "do people gender the voice" to "how do different users experience it", which has only been partially investigated in previous studies on ambiguous voices [9].

3. **Bias reduction vs. identity inclusion** ***(Conceptual)***: A central contribution of this thesis is clarifying that ambiguous voices do not

automatically create inclusive experiences for queer and nonbinary people. Instead, ambiguous voices primarily support inclusivity by reducing gendering and disrupting normative assumptions. I show that identity-based inclusion requires community-defined labels and representations, which ambiguous voices cannot provide. This conceptual distinction is missing in existing literature, where inclusivity claims are often made without empirical grounding or without input from gender-diverse users [84, 85].

4. **New methods for voice-robot design *(Methodological)***: Papers III and IV introduce sketching and physical prototyping as new ways to study how voices influence the robots people imagine. Earlier work mainly asked participants to match voices with existing robots [26, 60], which captures perception but not how people translate that perception into a design. In contrast, the sketching and prototyping methods reveal how voice cues translate into concrete visual and physical choices. They expose implicit gender assumptions that rating and matching tasks cannot uncover. Together, these methods provide new tools for studying how voice shapes gender in HRI.

5. **Ambiguity reduces gendering but not defaults *(Empirical)***: Across Papers III and IV, I show that ambiguous voices consistently reduce explicit gender attribution and encourage designs that avoid clear gender markers. However, participants still tend to default to masculine features, an effect linked to cultural patterns and to the "male-by-default" tendency discussed in previous work [48, 49]. My research offers the first step, using both sketching and physical prototyping, in showing how ambiguous voices shift gender expectations without fully removing them.

6. **Resources and frameworks for the community *(Practical)***: In this thesis, I contribute a set of practical resources to support research on voices beyond the gender binary. These include clear terminology for non-gendered voices, a public repository of collected gender-ambiguous voice samples, a small exploratory acoustic analysis to aid comparison between voices, a multi-dimensional approach for evaluating ambiguous voices, a dataset of 180 robot sketches, and the RoMix prototyping toolkit for studying gender and anthropomorphism in robot design. Together, these resources support more transparent, comparable, and replicable research on voice design and embodiment in HCI and HRI.

4.5 Reflections and Methodological Lessons

Alongside shaping my future work, these four papers also made visible what I would do differently if I ran the same studies today. They taught me as much about my future research practice as they did about ambiguous voices themselves.

In Paper I, for example, I now realise that I could have reached out more broadly to authors who used non-gendered voices, even if their papers did not meet all the inclusion criteria of the review. Doing so might have allowed a richer acoustic analysis and a more complete understanding of the landscape of ambiguous voices.

In Paper II, my main regret is not asking participants, especially nonbinary ones, open-ended questions about why they evaluated the voices the way they did. Their quantitative ratings showed clear patterns, but I missed the opportunity to understand the reasons behind those patterns in their own words. From this, I learned the value of combining quantitative and qualitative methods when studying voices and identity.

A similar reflection applies to Paper III. Although I did ask participants whether they sketched the robot with a gender in mind, I relied on an open-ended question rather than a more structured scale. This meant that my interpretations of these answers might not have been 100% accurate, and a quantitative measure could have given a clearer picture of participants' intentions.

In both Paper III and Paper IV, I now recognise how valuable it would have been to intentionally include nonbinary participants in equal numbers to men and women. Although I included balanced numbers of women, men, and nonbinary participants in the validation study of Paper IV, I was unable to follow the same approach in the main experiment. This was primarily due to practical constraints: recruiting sufficient in-person participants across three gender groups, while also accounting for multiple experimental conditions, would have required a substantially larger sample than was feasible. Reflecting on this limitation has highlighted the importance of developing more intentional and inclusive recruitment strategies, which I aim to prioritise in my future research.

Noticing these gaps does not diminish the value of the studies. Instead, they highlight how much my thinking has changed over these two years and how much space there is for improvement in future work.

4.6 Ideas for Future Work

The studies presented in this thesis, from a systematic review to online experiments and an in-person study, have provided me with some answers to my three research questions and, equally importantly, reshaped what comes next. Figure 4.9 summarises the path I followed in these two years: starting with a review of gender-ambiguous voices in HCI (Paper I), moving to how gender-ambiguous voices are perceived across women, men, and nonbinary participants (Paper II), how such voices influence mental images of robots (Paper III), and to how they shape the physical construction of robots (Paper IV). Together, these papers led me to the point represented at the end of Figure 4.9: **and now**?

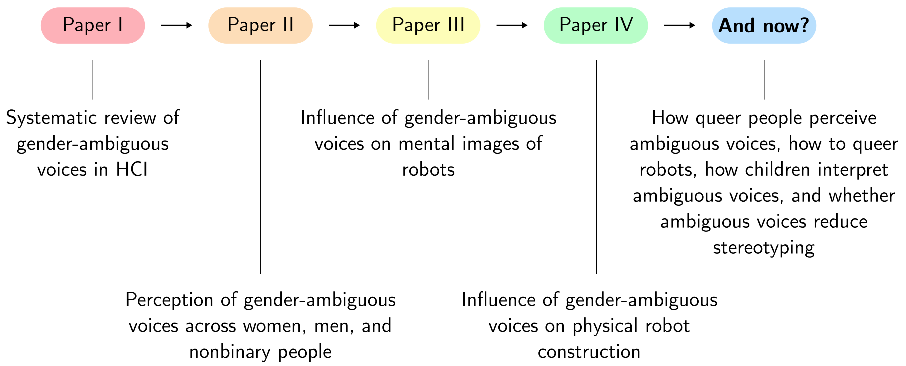

Figure 4.9: Summary of the research trajectory across the four papers and the emerging questions that shape my future work.

A moment that shaped the "and now", came when I presented my systematic review (Paper I) at the CHI conference on Human Factors in Computing Systems. During the Q&A, someone asked why I call these voices "ambiguous", because to them *"they just sounded like queer people"*. This question stayed with me. It made me rethink how I use labels in my work and later pushed me to write a currently under-review paper on how certain labels can misrepresent queer and nonbinary people. In that paper, I argue that if the aim is to be inclusive, we should consider using labels like "queer" and "nonbinary" rather than relying on terms such as "ambiguous" to describe non-gendered voices. It also opened a new direction: instead of only studying how people perceive ambiguous voices, I began thinking about what it could mean to intentionally explore how agents and voice assistants might express queerness in ways that matter to the communities they represent or serve. To move toward this bigger goal, my first future work is to better understand how queer and nonbinary listeners themselves perceive ambiguous voices. Unlike Paper II, which relied on structured ratings of pre-existing voices, with this work I want to focus on community-informed, qualitative and participatory methods to unpack why certain voices feel affirming or alienating, how labels shape perception, and what vocal qualities queer and nonbinary listeners associate with respectful representation.

Before continuing with my future work, it is important to clarify that the ideas proposed here form part of a larger plan. I am aware that I will not be able to address all of them within the remaining time of my PhD. These directions are not a to-do list, but starting points for discussion and for receiving input from you, the reader, as well as from my supervisors and my licentiate opponent on where this work could meaningfully go next.

4.6.1 What Do Nonbinary People Think of Ambiguous Voices?

One important insight I gained from Paper II is that gender-ambiguous voices do not resonate equally across listeners. Nonbinary participants evaluated these voices more critically than women and men. This finding was meaningful because it exposed a gap between what researchers often assume to be inclusive and what people with lived experience actually feel. I also realised that I had treated ambiguity as naturally inclusive, and I also thought that removing gender cues would automatically lead to fairer or more inclusive interactions. This study made it clear that this assumption does not hold for everyone and that the "ambiguous" category/label may be experienced as flattening, reductive, or misaligned with queer and nonbinary identities. This pushed me to think more critically about who gets to define inclusivity in design and whose perspectives are missing from these definitions, as explained before in section 4.2.

Because of this, I see a strong need to centre nonbinary and gender-expansive participants in my future research, not only as a demographic variable but as active collaborators. Their perspectives are essential for understanding the emotional, cultural, and political dimensions of gender-inclusive voices, which are often overlooked when design decisions are made solely through technical or aesthetic choices. In my future work, I want to use participatory and community-informed methods where nonbinary people describe, design, and evaluate voices that feel representative to them. This will help me understand what inclusion sounds like from their perspectives rather than imposing predefined labels, and it will support a more grounded understanding of how voice can express identities beyond the binary in ways that feel authentic, respectful, and meaningful to the communities they aim to represent.

4.6.2 How Can We Queer Robots?

The reflections above naturally led me to a broader question: if ambiguous voices alone cannot guarantee inclusion for queer and nonbinary people, then what would it mean for robots themselves to be designed with queer perspectives in mind? Recent work in robotics highlights how little the HRI field has engaged with queer perspectives, both in terms of representation within the research community and in the design of robotic systems themselves. Korpan et al. [85] point out that most robots reproduce binary gender norms through body shape, behaviour, and voice, and that queer people are rarely considered as stakeholders, even though robots inevitably interact with them. This lack of queer-informed design risks reproducing bias, exclusion, or misrepresentation. Integrating queer frameworks into robotics opens the possibility of robots whose identities, embodiments, and interaction styles resist normative assumptions rather than reinforcing them.

For this reason, as future work, I ask myself: *what would it actually mean to make a robot queer?* This is not something I imagined at the beginning of my PhD, but it has become an exciting direction for my future research. I want to understand what queering a robot actually means and what forms it could take. As Axelsson argues [83], queering a robot involves considering how queerness might be expressed through a robot's voice, body, movement, colour, behaviour, or overall interaction style, and recognising that queerness is not one single aesthetic or attitude but a wide range of expressions and possibilities.

To explore this, I plan to work closely with queer and nonbinary communities while experimenting with different combinations of robot features. I want to "play around" with designs, try less common ideas, and test expressions that do not follow typical masculine-feminine expectations. This may involve mixing voice cues in new ways, combining unexpected body shapes or materials, or creating movements that challenge standard interaction patterns. At the same time, I am interested in queerness as something flexible rather than fixed. Many queer and nonbinary people mix and match elements in their own self-expression, and I want to see whether robots can support this kind of expression too. By letting participants change a robot's voice, appearance, or behaviour, queer expression could emerge through personal choice rather than being predetermined by designers.

Recent work in queer robotics provides a foundation for this direction. Axelsson [83] shows how speculative design can help imagine queer robot futures, for example, by presenting robots that reflect queer in-group identities, robots that engage in queer activism, and networks of queer-owned robots that support community needs. These ideas illustrate that queering robots is not only about aesthetics, but also about imagining different social roles, relationships, and forms of care. At the same time, this line of work also draws attention to the limits and risks of such approaches. For example, Stolp-Smith and Williams [86] show how trans and nonbinary users perceive agender robot designs: while participants see potential for normalising non-cis identities, they also identify risks of caricature, backlash, and dehumanisation. Taken together, these perspectives underscore the need to approach queer robot design through inclusive, community-informed methods that prioritise authenticity and care.

These perspectives suggest another clear next step for my research: creating a space where queer researchers and practitioners can explore together what it means to queer robots. As future work, I would like to run a workshop at the HRI conference where queer researchers try out speculative design ideas, discuss ethical issues, and work together to outline guidelines for queer-informed robot design. This workshop would help turn theoretical ideas into practical directions and make queering robots a shared, community-led effort.

Overall, this future work aims to explore queerness as an intentional design practice. It builds on everything I learned during my first two years and opens a space for robots to express identities that are more diverse and more aligned with communities that have often been left out of HRI design [83, 85, 86].

4.6.3 How Do Children Perceive Ambiguous Voices?

As I think about queering robots and working with queer communities, another group has become increasingly important to me: children. This interest grew after I reviewed a currently unpublished paper discussing how vocal pitch influences children's perceptions, particularly in relation to stress regulation. Since ambiguous voices are, among many other things, created through pitch manipulation, that paper prompted a new question for me: *How do children perceive ambiguous voices?*

While writing the literature review and analysing my own data, I noticed that almost nothing is known about how children interpret gender-ambiguous voices. Most research on gender in technology focuses on adults, yet many children interact daily with voice assistants, conversational agents, and robots at home and in educational settings.

I believe that investigating ambiguous voices with children will be both valuable and challenging. Studies with children come with specific methodological and ethical considerations, which I am eager to develop, but understanding how they reason about gendered and non-gendered voices is essential for designing technologies that do not unintentionally reinforce stereotypes from an early age. Understanding children's reactions is especially important because early interactions with technology shape how gender stereotypes form or shift [87].

In future work, I want to explore this through "child-friendly" methods such as sketching, storytelling, and playful robot interactions. For example, children could interact with a small robot or voice assistant over several short sessions, where the only thing that changes between groups is the voice (ambiguous, masculine, or feminine). After each session, they could be asked to tell a short story about "who the robot is" and "what it does". Comparing these stories across conditions and age groups would reveal how ambiguity interacts with early gender reasoning and whether children respond differently from adults. This work would not only fill a gap in the literature but also connect back to the broader question of what it means to design inclusive technologies for all users, not only adults and queers.

4.6.4 Can a Gender-Ambiguous Voice Reduce Stereotyping?

And now we reach the question that will likely take the most time to investigate. Thinking about children also led me to a broader limitation in my studies: almost everything I know about gender-ambiguous voices comes from first impressions. Yet voice assistants and robots are not one-time encounters. Long-term HRI research shows that repeated interactions can change perception, engagement, and even users' relationships with a robot. Matheus et al. [88] show that many social effects, such as trust, comfort, and perceived robot qualities, shift only after the initial interaction. In other words, single-session experiments cannot capture the dynamics that unfold over days or weeks. Longitudinal studies also demonstrate that people's feelings about a robot often change over time, for example, becoming more comfortable or forming more stable impressions. This is also highlighted in a recent review of gendering in HRI by Perugia and Lisy [45]. They argue that stereotype attribution in HRI is typically studied using static stimuli and short-term designs, and they call for work that examines how gendered impressions and stereotypes evolve.

These observations raise important questions for my work: *Can a gender-ambiguous voice reduce stereotyping over time?* In particular, *What happens when ambiguity is experienced repeatedly rather than briefly?*, and *Does a gender-ambiguous voice become more familiar and less gendered with continued use, or do users continue to map it back onto binary categories?* I believe these questions can only be meaningfully addressed through a longitudinal study. One possible approach is to expose participants to a voice assistant or robot with an ambiguous, masculine, or feminine voice across multiple sessions, allowing comparison of how perceptions evolve over time and whether gendering becomes weaker or stronger across conditions.

Such evidence would be highly valuable to the HRI community, where longitudinal studies remain comparatively rare, and there is limited understanding of how voice impressions evolve, particularly for ambiguous voices. Building on the methodological foundation established in Papers I–IV, this approach is feasible within the remaining time of this PhD and would help clarify whether gender-ambiguous voices support less gendered and less stereotyped interactions in everyday use rather than only in short-term experimental settings.

4.6.5 To Summarise

Overall, these directions point toward a broader research plan that moves beyond categorising voices and toward understanding how gender, identity, and technology intersect in everyday life. My future work aims to build a more holistic picture of how users make sense of gendered cues in HRI by examining queer and nonbinary experiences, exploring what it means to queer robots, investigating how children develop gendered expectations in relation to technology, and testing whether ambiguity can reduce stereotyping over time. This broader perspective will help me move from treating ambiguity as a single design choice to seeing it as part of a larger process that shapes how people relate to agents, how identities are recognised, and how inclusion unfolds in practice.

Bibliography

[1] N. S. Eidsheim and K. Meizel, *The Oxford Handbook of Voice Studies*. Oxford University Press, Jul. 2019, ISBN: 9780199982295. DOI: 10.1093/oxfordhb/9780199982295.001.0001. [Online]. Available: https://doi.org/10.1093/oxfordhb/9780199982295.001.0001 (cit. on pp. 3, 4).

[2] J. W. Mullennix, S. E. Stern, S. J. Wilson and C. lynn Dyson, "Social perception of male and female computer synthesized speech," *Computers in Human Behavior*, vol. 19, no. 4, pp. 407–424, 2003, ISSN: 0747-5632. DOI: https://doi.org/10.1016/S0747-5632(02)00081-X (cit. on p. 3).

[3] A. Mari, A. Mandelli and R. Algesheimer, "Empathic voice assistants: Enhancing consumer responses in voice commerce," *Journal of Business Research*, vol. 175, p. 114 566, 2024, ISSN: 0148-2963. DOI: https://doi.org/10.1016/j.jbusres.2024.114566. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0148296324000705 (cit. on p. 4).

[4] F. Eyssel and F. Hegel, "(s)he's got the look: Gender stereotyping of robots 1," *Journal of Applied Social Psychology*, vol. 42, Jul. 2012. DOI: 10.1111/j.1559-1816.2012.00937.x (cit. on pp. 4, 12, 16).

[5] C. Crowell, M. Scheutz, P. Schermerhorn and M. Villano, "Gendered voice and robot entities: Perceptions and reactions of male and female subjects," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, New York, NY, USA: IEEE Press, Oct. 2009, pp. 3735–3741 (cit. on pp. 4, 12).

[6] A. H.-C. Hwang and A. S. Won, "The sound of support: Gendered voice agent as support to minority teammates in gender-imbalanced team," in *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, ser. CHI '24, New York, NY, USA: Association for Computing Machinery, 2024, ISBN: 9798400703300. DOI: 10.1145/3613904.3642202. [Online]. Available: https://doi.org/10.1145/3613904.3642202 (cit. on p. 4).

[7] S. Tolmeijer, N. Zierau, A. Janson, J. S. Wahdatehagh, J. M. M. Leimeister and A. Bernstein, "Female by default? exploring the effect of voice assistant gender and pitch on trait and trust attribution," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*,

ser. CHI EA '21, New York, NY, USA: Association for Computing Machinery, 2021, ISBN: 9781450380959. DOI: 10.1145/3411763.3451623. [Online]. Available: https://doi.org/10.1145/3411763.3451623 (cit. on pp. 4–6, 28).

[8] S. Song, J. Baba, J. Nakanishi, Y. Yoshikawa and H. Ishiguro, "Mind the voice!: Effect of robot voice pitch, robot voice gender, and user gender on user perception of teleoperated robots," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '20, New York, NY, USA: Association for Computing Machinery, 2020, 18, ISBN: 9781450368193. DOI: 10.1145/3334480.3382988. [Online]. Available: https://doi.org/10.1145/3334480.3382988 (cit. on p. 4).

[9] A. Danielescu, S. A. Horowit-Hendler, A. Pabst, K. M. Stewart, E. M. Gallo and M. P. Aylett, "Creating inclusive voices for the 21st century: A non-binary text-to-speech for conversational assistants," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23, New York, NY, USA: Association for Computing Machinery, 2023, ISBN: 9781450394215. DOI: 10.1145/3544548.3581281. [Online]. Available: https://doi.org/10.1145/3544548.3581281 (cit. on pp. 4, 17, 19, 23, 24, 28–30, 37).

[10] American Psychological Association. "Cisgender." Accessed: 2025-11-19, APA Dictionary of Psychology. [Online]. Available: https://dictionary.apa.org/cisgender (cit. on p. 4).

[11] K. Ito, *The lj speech dataset*, https://keithito.com/LJ-Speech-Dataset/, 2017 (cit. on p. 4).

[12] R. Zandie, M. H. Mahoor, J. Madsen and E. S. Emamian, "Ryanspeech: A corpus for conversational text-to-speech synthesis," in *Interspeech*, 2021 (cit. on p. 4).

[13] S. T. Längle, S. Schlögl, A. Ecker, W. S. M. T. van Kooten and T. Spieß, "Nonbinary voices for digital assistantsan investigation of user perceptions and gender stereotypes," *Robotics*, vol. 13, no. 8,, 2024, ISSN: 2218-6581. DOI: 10.3390/robotics13080111. [Online]. Available: https://www.mdpi.com/2218-6581/13/8/111 (cit. on pp. 4, 18, 19).

[14] I. Torre, E. Lagerstedt, N. Dennler, K. Seaborn, I. Leite and E. Székely, "Can a gender-ambiguous voice reduce gender stereotypes in human-robot interactions?" In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, New York, NY, USA: IEEE Press, 2023, pp. 106–112. DOI: 10.1109/RO-MAN57019.2023.10309500 (cit. on pp. 4, 5, 18, 25, 28).

[15] K. Winkle, E. Lagerstedt, I. Torre and A. Offenwanger, "15 years of (who)man robot interaction: Reviewing the h in human-robot interaction," *J. Hum.-Robot Interact.*, vol. 12, no. 3, Apr. 2023. DOI: 10.1145/3571718. [Online]. Available: https://doi.org/10.1145/3571718 (cit. on p. 4).

[16] A. Offenwanger, A. J. Milligan, M. Chang, J. Bullard and D. Yoon, "Diagnosing bias in the gender representation of hci research participants: How it happens and where we are," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21, New York, NY, USA: Association for Computing Machinery, 2021, ISBN: 9781450380966. DOI: 10.1145/3411764.3445383. [Online]. Available: https://doi.org/10.1145/3411764.3445383 (cit. on p. 4).

[17] K. Spiel, O. L. Haimson and D. M. Lottridge, "How to do better with gender on surveys: A guide for hci researchers," *Interactions*, vol. 26, no. 4, pp. 62–65, 2019. DOI: 10.1145/3338283. [Online]. Available: https://doi.org/10.1145/3338283 (cit. on p. 4).

[18] K. Spiel, O. Keyes and P. Barlas, "Patching gender: Non-binary utopias in hci," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '19, New York, NY, USA: Association for Computing Machinery, 2019, 111, ISBN: 9781450359719. DOI: 10.1145/3290607.3310425. [Online]. Available: https://doi.org/10.1145/3290607.3310425 (cit. on pp. 4, 18).

[19] M. De Cet, M. Obaid and I. Torre, "Breaking the binary: A systematic review of gender-ambiguous voices in human-computer interaction," in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, ser. CHI '25, New York, NY, USA: Association for Computing Machinery, 2025, ISBN: 9798400713941. DOI: 10.1145/3706598.3713608. [Online]. Available: https://doi.org/10.1145/3706598.3713608 (cit. on pp. 5, 22, 28).

[20] J. Butler, *Gender Trouble: Feminism and the Subversion of Identity.* New YOrk: Routledge, 1990, ISBN: 9780415389556 (cit. on pp. 5, 15).

[21] G. Rubin, "The traffic in women: Notes on the political economy of sex," in *Toward an Anthropology of Women*, R. R. Reiter, Ed., New York: Monthly Review Press, 1975, pp. 157–210 (cit. on pp. 5, 15).

[22] S. J. Sutton, "Gender ambiguous, not genderless: Designing gender in voice user interfaces (vuis) with sensitivity," in *Proceedings of the 2nd Conference on Conversational User Interfaces*, ser. CUI '20, New York, NY, USA: Association for Computing Machinery, 2020, ISBN: 9781450375443. DOI: 10.1145/3405755.3406123. [Online]. Available: https://doi.org/10.1145/3405755.3406123 (cit. on pp. 5, 6, 16, 18, 37).

[23] I. Jestin, J. Fischer, M. J. Galvez Trigo, D. Large and G. Burnett, "Effects of wording and gendered voices on acceptability of voice assistants in future autonomous vehicles," in *Proceedings of the 4th Conference on Conversational User Interfaces*, ser. CUI '22, New York, NY, USA: Association for Computing Machinery, 2022, ISBN: 9781450397391. DOI: 10.1145/3543829.3543836. [Online]. Available: https://doi.org/10.1145/3543829.3543836 (cit. on pp. 5, 28).

[24] C. Yu, C. Fu, R. Chen and A. Tapus, "First attempt of gender-free speech style transfer for genderless robot," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, New York, NY, USA: IEEE Press, 2022, pp. 1110–1113. DOI: 10.1109/HRI53351.2022.9889533 (cit. on pp. 5, 18).

[25] S. J. Sutton, P. Foulkes, D. Kirk and S. Lawson, "Voice as a design material: Sociophonetic inspired design strategies in human-computer interaction," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19, New York, NY, USA: Association for Computing Machinery, 2019, 114, ISBN: 9781450359702. DOI: 10.1145/3290605.3300833. [Online]. Available: https://doi.org/10.1145/3290605.3300833 (cit. on p. 6).

[26] M. De Cet, M. Cvajner, I. Torre and M. Obaid, "Do your expectations match? a mixed-methods study on the association between a robots voice and appearance," *ACM Conversational User Interfaces*, 11 pages, 2024. [Online]. Available: https://doi.org/10.1145/3640794.3665551 (cit. on pp. 6, 16, 38).

[27] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati and F. Tanaka, "Social robots for education: A review," *Science Robotics*, vol. 3, eaat5954, Aug. 2018. DOI: 10.1126/scirobotics.aat5954 (cit. on p. 10).

[28] M. Alemi, A. Meghdari and M. Ghazisaedy, "Employing humanoid robots for teaching english language in iranian junior high-schools," *Int. J. Humanoid Robot.*, vol. 11, no. 03, p. 1 450 022, 2014. DOI: 10.1142/S0219843614500224 (cit. on p. 10).

[29] N. Rakhymbayeva, A. Amirova and A. Sandygulova, "A long-term engagement with a social robot for autism therapy," *Front. Robot. AI*, vol. 8, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frobt.2021.669972 (cit. on p. 10).

[30] A. Darmudas, E. Sobolewska and C. Strathearn, "Social robots in education: A biometrically driven approach to supporting wellbeing," in *Proceedings of the 38th International BCS Human-Computer Interaction Conference*, ser. BCS HCI '25, Swindon, GBR: BCS Learning & Development Ltd, 2025, 272283. DOI: 10.14236/ewic/BCSHCI2025.26. [Online]. Available: https://doi.org/10.14236/ewic/BCSHCI2025.26 (cit. on p. 10).

[31] K. Bussey and A. Bandura, "Social cognitive theory of gender development and differentiation," *Psychological Review*, vol. 106, no. 4, pp. 676–713, 1999. DOI: 10.1037/0033-295X.106.4.676 (cit. on p. 11).

[32] M. Harper and W. J. Schoeman, "Influences of gender as a basic-level category in person perception on the gender belief system," *Sex Roles*, vol. 49, pp. 517–526, 2003. DOI: 10.1023/A:1025884723841 (cit. on p. 11).

[33] J. S. Hyde, R. S. Bigler, D. Joel, C. C. Tate and S. M. van Anders, "The future of sex and gender in psychology: Five challenges to the gender binary," *American Psychologist*, vol. 74, no. 2, pp. 171–193, 2019, Place: US Publisher: American Psychological Association, ISSN: 1935-990X(Electronic),0003-066X(Print). DOI: 10.1037/amp0000307 (cit. on p. 11).

[34] G. Perugia, S. Guidi, M. Bicchi and O. Parlangeli, "The shape of our bias: Perceived age and gender in the humanoid robots of the abot database," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '22, New York, NY, USA: IEEE Press, 2022, 110119 (cit. on pp. 11, 13).

[35] C. L. Martin and D. Ruble, "Childrens search for gender cues: Cognitive perspectives on gender development," *Current Directions in Psychological Science*, vol. 13, no. 2, 6770, Apr. 2004, ISSN: 1467-8721. DOI: 10.1111/j.0963-7214.2004.00276.x. [Online]. Available: http://dx.doi.org/10.1111/j.0963-7214.2004.00276.x (cit. on p. 11).

[36] J. Urakami and K. Seaborn, "Nonverbal cues in humanrobot interaction: A communication studies perspective," *ACM Transactions on Human-Robot Interaction*, vol. 12, no. 2, 121, Mar. 2023, ISSN: 2573-9522. DOI: 10.1145/3570169. [Online]. Available: http://dx.doi.org/10.1145/3570169 (cit. on p. 11).

[37] J. Robertson, "Gendering humanoid robots: Robo-sexism in japan," *Body and Society - BODY SOC*, vol. 16, pp. 1–36, Jul. 2010. DOI: 10.1177/1357034X10364767 (cit. on p. 11).

[38] M. Siegel, C. Breazeal and M. I. Norton, "Persuasive robotics: The influence of robot gender on human behavior," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, New York, NY, USA: IEEE, 2009, pp. 2563–2568. DOI: 10.1109/IROS.2009.5354116 (cit. on p. 11).

[39] R. B. Cialdini, *Influence: The Psychology of Persuasion*, Revised edition. New York: Quill/William Morrow, 1993 (cit. on p. 11).

[40] F. Eyssel, D. Kuchenbrandt, S. Bobinger, L. de Ruiter and F. Hegel, "'if you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism," in *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '12, New York, NY, USA: Association for Computing Machinery, 2012, 125126, ISBN: 9781450310635. DOI: 10.1145/2157689.2157717. [Online]. Available: https://doi.org/10.1145/2157689.2157717 (cit. on pp. 11, 16).

[41] G. Aşkın, İ. Saltık, T. E. Boz and B. A. Urgen, "Gendered actions with a genderless robot: Gender attribution to humanoid robots in action," *International Journal of Social Robotics*, vol. 15, no. 11, pp. 1915–1931, 2023 (cit. on pp. 12, 16).

[42] P. Bisconti and G. Perugia, "How do we gender robots?: Inquiring the relationship between perceptual cues and context of use," English, in *GenR 2021 Workshop on Gendering Robots*, GenR Workshop Gendering Robots, 2021 : Ongoing (Re)configurations of Gender in Robotics, RO-MAN 2021 ; Conference date: 08-08-2021 Through 12-11-2021, Aug. 2021. [Online]. Available: `https://easychair.org/cfp/GenR2021` (cit. on p. 12).

[43] B. Tay, Y. Jung and T. Park, "When stereotypes meet robots: The double-edge sword of robot gender and personality in human–robot interaction," *Computers in Human Behavior*, vol. 38, pp. 75–84, 2014 (cit. on p. 12).

[44] M. Fridin and M. Belokopytov, "Acceptance of socially assistive humanoid robot by preschool and elementary school teachers," *Comput. Hum. Behav.*, vol. 33, 2331, 2014, ISSN: 0747-5632. DOI: `10.1016/j.chb.2013.12.016`. [Online]. Available: `https://doi.org/10.1016/j.chb.2013.12.016` (cit. on p. 12).

[45] G. Perugia and D. Lisy, "Robots gendering trouble: A scoping review of gendering humanoid robots and its effects on hri," *International Journal of Social Robotics*, vol. 15, pp. 1725–1753, 2023. DOI: `10.1007/s12369-023-01061-6`. [Online]. Available: `https://link.springer.com/article/10.1007/s12369-023-01061-6` (cit. on pp. 12, 16, 44).

[46] D. J. Haraway, *Modest Witness@Second Millennium. FemaleMan Meets OncoMouse: Feminism and Technoscience*. New York and London: Routledge, 1997, ISBN: 0-415-91245-8 (cit. on p. 13).

[47] H. Han, Y. Lee, C. Zhang, J. Lu and L. Wang, "He or she? a male-default bias in chatbot gender attribution across explicit and implicit measures," *Computers in Human Behavior Reports*, p. 100 860, 2025, ISSN: 2451-9588. DOI: `https://doi.org/10.1016/j.chbr.2025.100860`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S2451958825002751` (cit. on p. 13).

[48] J. Bernotat, F. Eyssel and J. Sachse, "The (fe) male robot: How robot body shape impacts first impressions and trust towards robots," *International Journal of Social Robotics*, vol. 13, no. 3, pp. 477–489, 2021 (cit. on pp. 13, 25, 38).

[49] I. Hipólito, K. Winkle and M. Lie, "Enactive artificial intelligence: Subverting gender norms in human-robot interaction," *Frontiers in Neurorobotics*, vol. 17, p. 1 149 303, 2023 (cit. on pp. 13, 38).

[50] interACT: Advocates for Intersex Youth, *Intersex definitions*, Accessed: 2025-08-20, 2025. [Online]. Available: `https://interactadvocates.org/intersex-definitions/` (cit. on p. 15).

[51] A. Fausto-Sterling, "Dueling dualisms," in *Sexing the Body: Gender Politics and the Construction of Sexuality*, New York: Perseus Books Group, 2000 (cit. on p. 15).

[52] J. Butler, *Bodies That Matter: On the Discursive Limits of Sex*. New York: Routledge, 1993 (cit. on p. 15).

[53] E. I. for Gender Equality, *Gender-neutral*, Accessed: 2024-12-04, 2024. [Online]. Available: https://eige.europa.eu/publications-resources/thesaurus/terms/1321?language_content_entity=en (cit. on pp. 16, 37).

[54] J. Cambre and C. Kulkarni, "One voice fits all? social implications and research challenges of designing voices for smart devices," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, no. CSCW, Nov. 2019. DOI: 10.1145/3359325. [Online]. Available: https://doi.org/10.1145/3359325 (cit. on p. 16).

[55] S. F. Warta, "If a robot did the robot, would it still be called the robot or just dancing? perceptual and social factors in human-robot interactions," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, no. 1, pp. 796–800, 2015. DOI: 10.1177/1541931215591244. [Online]. Available: https://doi.org/10.1177/1541931215591244 (cit. on p. 16).

[56] J. Otterbacher and M. Talias, "S/he's too warm/agentic! the influence of gender on uncanny reactions to robots," in *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*, 2017, pp. 214–223 (cit. on p. 16).

[57] M. K. Strait, C. Aguillon, V. Contreras and N. Garcia, "The public's perception of humanlike robots: Online social commentary reflects an appearance-based uncanny valley, a general fear of a technology takeover, and the unabashed sexualization of female-gendered robots," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2017, pp. 1418–1423. DOI: 10.1109/ROMAN.2017.8172490 (cit. on p. 16).

[58] K. Seaborn, N. P. Miyake, P. Pennefather and M. Otake-Matsuura, "Voice in human–agent interaction: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 4, pp. 1–43, 2021. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3386867 (cit. on p. 16).

[59] K. Seaborn and P. Pennefather, "Neither "hear" nor "their": Interrogating gender neutrality in robots," Mar. 2022. DOI: 10.1109/HRI53351.2022.9889350 (cit. on p. 16).

[60] C. McGinn and I. Torre, "Can you tell the robot by the voice? an exploratory study on the role of voice in the perception of robots," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 211–221. DOI: 10.1109/HRI.2019.8673305 (cit. on pp. 16, 18, 38).

[61] Speechgen.io, *Speechgen.io*, Accessed: 2024-05-18, 2024. [Online]. Available: https://speechgen.io (cit. on p. 17).

[62] CereProc, *Cereproc*, Accessed: 2024-05-18, 2024. [Online]. Available: https://www.cereproc.com (cit. on p. 17).

[63] E. Székely and M. Hope, "An inclusive approach to creating a palette of synthetic voices for gender diversity," in *Interspeech 2024*, 2024, pp. 3070–3074. DOI: 10.21437/Interspeech.2024-1543 (cit. on pp. 17, 19).

[64] C. Callaway and K. Simaan, "Wired for speech: How voice activates and advances the human-computer relationship," *Comput. Linguist.*, vol. 32, pp. 451–452, 2006. DOI: 10.1162/coli.2006.32.3.451 (cit. on p. 17).

[65] D. Riding, D. Lonsdale and B. Brown, "The effects of average fundamental frequency and variance of fundamental frequency on male vocal attractiveness to women," *J. Nonverbal Behav.*, vol. 30, no. 2, pp. 55–61, 2006. DOI: 10.1007/s10919-006-0005-3 (cit. on p. 17).

[66] B. C. Jones, D. R. Feinberg, L. M. DeBruine, A. C. Little and J. Vukovic, "Integrating cues of social interest and voice pitch in mens preferences for womens voices," *Biol. Lett.*, vol. 4, no. 2, pp. 192–194, 2008. DOI: 10.1098/rsbl.2007.0626 (cit. on p. 17).

[67] S. A. Collins and C. Missing, "Vocal and visual attractiveness are related in women," *Anim. Behav.*, vol. 65, no. 5, pp. 997–1004, 2003. DOI: 10.1006/anbe.2003.2123 (cit. on p. 17).

[68] I. Torre, L. White, J. Goslin and S. Knight, "The irrepressible influence of vocal stereotypes on trust," *Quarterly Journal of Experimental Psychology*, vol. 77, no. 10, pp. 1957–1966, 2024 (cit. on p. 17).

[69] O. Niebuhr, J. Voße and A. Brem, "What makes a charismatic speaker? a computer-based acoustic-prosodic analysis of steve jobs tone of voice," *Computers in Human Behavior*, vol. 64, pp. 366–382, 2016 (cit. on p. 17).

[70] A. Niculescu, B. van Dijk, A. Nijholt and S. L. See, "The influence of voice pitch on the evaluation of a social robot receptionist," in *2011 International Conference on User Science and Engineering (i-USEr)*, 2011, pp. 18–23. DOI: 10.1109/iUSEr.2011.6150529 (cit. on p. 17).

[71] D. Becker et al., "Influence of robots voice naturalness on trust and compliance," *J. Hum.-Robot Interact.*, vol. 14, no. 2, Jan. 2025. DOI: 10.1145/3706066. [Online]. Available: https://doi.org/10.1145/3706066 (cit. on p. 17).

[72] M. Mara, S. Schreibelmayr and F. Berger, "Hearing a nose?: User expectations of robot appearance induced by different robot voices," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2020, 355356. DOI: 10.1145/3371382.3378285 (cit. on p. 18).

[73] I. Torre, A. B. Latupeirissa and C. McGinn, "How context shapes the appropriateness of a robots voice," in *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2020, 215222. DOI: 10.1109/RO-MAN47096.2020.9223449 (cit. on p. 18).

[74] S. Mooshammer and K. Etzrodt, "Social research with gender-neutral voices in chatbots the generation and evaluation of artificial gender-neutral voices with praat and google wavenet," in Jan. 2022, pp. 176–191, ISBN: 978-3-030-94889-4. DOI: 10.1007/978-3-030-94890-0_11 (cit. on p. 18).

[75] E. Székely, J. Gustafson and I. Torre, "Prosody-controllable Gender-ambiguous Speech Synthesis: A Tool for Investigating Implicit Bias in Speech Perception," in *Proc. INTERSPEECH 2023*, 2023, pp. 1234–1238. DOI: 10.21437/Interspeech.2023-2086 (cit. on p. 18).

[76] M. De Cet, K. Seaborn, M. Obaid and I. Torre, "Hearing ambiguity: Exploring beyond-gender impressions of artificial ambiguous voices," in *Proceedings of the 7th ACM Conference on Conversational User Interfaces*, ser. CUI '25, New York, NY, USA: Association for Computing Machinery, 2025, ISBN: 9798400715273. DOI: 10.1145/3719160.3736622. [Online]. Available: https://doi.org/10.1145/3719160.3736622 (cit. on pp. 23, 28, 29, 31).

[77] K. Seaborn, M. Altmeyer, G.. Li, B. Ku, S. Kobuki and J. Urakami, "The voice experience inventory (voxi): Validating a consensus-driven instrument for measuring user impressions of computer voice," *International Journal of Human-Computer Studies*, vol. 203, p. 103 576, 2025, ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2025.103576. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581925001338 (cit. on pp. 23, 29, 37).

[78] M. De Cet, M. Sturdee, M. Obaid and I. Torre, "Sketching robots: Exploring the influence of gender-ambiguous voices on robot perception," in *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '25, IEEE Press, 2025, 103112. DOI: 10.5555/3721488.3721505. [Online]. Available: https://dl.acm.org/doi/10.5555/3721488.3721505 (cit. on pp. 24, 28, 29, 33, 35).

[79] M. De Cet, N. Hashmati, M. Obaid and I. Torre, "From voice to form: How gender-ambiguous voices shape physical robot design," in *Proceedings of the 2026 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '26, New York, NY, USA: Association for Computing Machinery, 2026. DOI: 10.1145/3757279.3785559 (cit. on pp. 25, 35).

[80] C. G. Christiansen, S. Hardt, S. F. Jensen, K. Fischer and O. Palinko, "Speech impact in a usability test - a case study of the kubo robot," in *2022 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, New York, NY, USA: IEEE Press, 2022, pp. 723–726. DOI: 10.1109/HRI53351.2022.9889621 (cit. on p. 28).

[81] I. Lopatovska, D. Brown and E. Korshakova, "Contextual perceptions of feminine-, masculine- and gender-ambiguous-sounding conversational agents," in *Information for a Better World: Shaping the Global Future*, M. Smits, Ed., Cham: Springer International Publishing, 2022, pp. 459–480 (cit. on p. 28).

[82] J. M. Kuch, F. Melchior and C. Becker-Asano, "Effects of gender neutralization on the anthropomorphism of natural and synthetic voices," in *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, New York, NY, USA: IEEE Press, 2023, pp. 2080–2085. DOI: 10.1109/RO-MAN57019.2023.10309479 (cit. on p. 28).

[83] M. Axelsson, *Speculative design of equitable robotics: Queer fictions and futures*, 2025. arXiv: 2509.01643 [cs.RO]. [Online]. Available: https://arxiv.org/abs/2509.01643 (cit. on pp. 32, 42, 43).

[84] M. Hope, "Creation, perception, and use of gender expansive synthetic voices," Ph.D. dissertation, University of Delaware, Newark, DE, 2024. DOI: 10.58088/5733-z873. [Online]. Available: https://udspace.udel.edu/handle/19716/35413 (cit. on pp. 32, 38).

[85] R. Korpan et al., "Launching queer in robotics [women in engineering]," *IEEE Robotics and Automation Magazine*, vol. 31, no. 2, pp. 144–146, 2024. DOI: 10.1109/MRA.2024.3388277 (cit. on pp. 38, 41, 43).

[86] M. Stolp-Smith and T. Williams, "More than binary: Transgender and non-binary perspectives on human robot interaction," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '24, New York, NY, USA: Association for Computing Machinery, 2024, 697705, ISBN: 9798400703225. DOI: 10.1145/3610977.3634939. [Online]. Available: https://doi.org/10.1145/3610977.3634939 (cit. on pp. 42, 43).

[87] L. McGuire et al., "Stem gender stereotypes from early childhood through adolescence at informal science centers," *Journal of Applied Developmental Psychology*, vol. 67, p. 101 109, 2020. DOI: 10.1016/j.appdev.2020.101109. [Online]. Available: https://doi.org/10.1016/j.appdev.2020.101109 (cit. on p. 43).

[88] K. Matheus, R. Ramnauth, B. Scassellati and N. Salomons, "Long-term interactions with social robots: Trends, insights, and recommendations," *J. Hum.-Robot Interact.*, vol. 14, no. 3, Jun. 2025. DOI: 10.1145/3729539. [Online]. Available: https://doi.org/10.1145/3729539 (cit. on p. 44).