HUMAN-MACHINE COMMUNICATION

Human-Machine Communication
Special Edition, Volume 7, 2024
https://doi.org/10.30658/hmc.7.3

# A New Family Member or Just Another Digital Interface? Smart Speakers in the Lives of Families With Young Children

**Giovanna Mascheroni**[1]

1 Department of Communication, Università Cattolica del Sacro Cuore, Milano, Italy

## Abstract

Based on longitudinal qualitative research involving 20 families with at least one child aged 8 or younger, the article provides an account of how families, as distinctive communicative figurations, adopt, use, and make sense of smart speakers through diverse socially situated practices. Findings show that parents and children enter in a communicative relationship with smart speakers based on their attribution of human-like or machine-like traits to the device, and the device's response to their expectations. Moreover, engaging in communicative practices *through* and with smart speakers, family members subvert or reinforce existing power relations. However, smart speakers acquire new agency by intensifying the datafication and algorithmization of everyday life, thus entailing a shift in power dynamics between humans and machines.

**Keywords:** smart speakers, voice-based agents, children, families, deep mediatization

## Acknowledgments

This research was supported by Fondazione Cariplo (Bando Ricerca Sociale 2019).

## Introduction

While not initially designed for children—in fact, the first Echo Dot Kids Edition was only launched in May 2018—smart speakers have become part of children's everyday media

**CONTACT** Giovanna Mascheroni • giovanna.mascheroni@unicatt.it • Department of Communication • Università Cattolica del Sacro Cuore • Largo Gemelli 1 • 20123 Milano • Italy

ISSN 2638-602X (print)/ISSN 2638-6038 (online)
www.hmcjournal.com

repertoires in many countries (Rideout & Robb, 2020; Wald et al., 2023). For example, the proportion of US families with young children who owned at least a smart speaker rose from 9% in 2017 to 41% in 2020 (Rideout & Robb, 2020). Similarly, in 2020, 46% of Italian parents of young children owned one or more smart speakers (Zaffaroni et al., 2022). Yet, this rapid diffusion has only partially been matched by empirical research on the domestication of smart speakers, especially outside of the US. In fact, while most research to date has focused on usage practices (Lopatovska et al., 2019), or trust and privacy issues (Pridmore et al., 2019), knowledge of the meaning-making practices, negotiations, and conflicts around smart speakers in families with young children is still sparse.

Based on a longitudinal qualitative research involving 20 families with at least one child aged 8 or younger in Italy, this article aims to fill this knowledge gap by providing an account of the diverse socially situated practices and contexts in which families, as distinctive communicative figurations (Couldry & Hepp, 2017; Hepp & Hasebrink, 2018), adopt, use, and make sense of smart speakers—voice-based home devices with an integrated conversational agent. The notion of communicative figurations refers to a *constellation of actors*, who share a *frame of relevance* (a set of cultural values, norms, and orientations, including technological imaginaries), on which basis they interact through *communicative practices* supported by a distinctive *media ensemble*. In this light, each family's communication repertoire is grounded in their culture as much as in their media ensemble (Hepp & Hasebrink, 2018). Patterns of similarities and differences emerge across families based on the specific constellation of actors, their relevance frames, the distinctive set of communicative practices as well as the particular media ensemble.

Communicative figurations have been theorized as an analytical device to advance a non-mediacentric research on mediatization (Couldry & Hepp, 2017; Hepp, 2020a). The ongoing deep mediatization means that every sphere of the social, families included, has become mediatized, that is, increasingly interdependent on communicative practices and digital infrastructures. Mediatized families depend on "data relations" (Couldry & Mejias, 2019, p. 27): namely, on communicative practices, digital media, and technological infrastructures that extract, analyze, and communicate data, turning the home and families' everyday lives into data commodities (Mascheroni & Siibak, 2021). Indeed, we can understand datafication as both an outcome and an intensifier of mediatization (Mascheroni & Siibak, 2021)—an outcome of mediatization, because everyday mediatized practices generate an unprecedented volume of data traces; and, at the same time, an accelerator of mediatization, because, as the everyday opportunities for interaction and agency are increasingly defined by algorithmic-based automations, digital media are more and more embedded into the texture of users' everyday lives. In this context, we argue that a figurational approach to the domestication of smart speakers by families with young children is best suited to both contextualize these technologies within the longer mediatization of the home (Silverstone, 1999) and simultaneously account for the novelty of this specific communicative AI. In fact, the media have been constitutive of the modern idea, and experience, of the home as both an autonomous and relational entity: in this respect, smart speakers may be conceived of as the latest technological development in the process of "mobile privatization" (Williams, 2003 [1974], p. 19) and in the more recent mediatization of family life (Mascheroni & Siibak, 2021). However, as a specific form of communicative AI, smart speakers challenge our established definitions of media, and their role within mediatized domesticity, as we will further elaborate.

Furthermore, being essentially non-mediacentric, a figurational approach helps us dismantle the myth of communicative AI "as a completely autonomous technology" (Natale & Guzman, 2022, p. 629)—a myth that has often led to underplay human agency vis-à-vis machine agency. Conversely, and in line with a mediatization research agenda, we aim to understand how changes in the media ensemble following the introduction of smart speakers are associated with changing practices of communication based on each family's distinctive frames of relevance, and whether such changes lead to transformations in the power relations between family members. More specifically, we formulated the following research questions: To what extent does the family as a communicative figuration change in light of the domestication of smart speakers? What distinguishes the communicative practices *through* and *with* smart speakers from other mediated communicative practices within the family?

## Research on Children, Families, and Smart Speakers

Since conversational agents, such as those embedded in smart speakers, democratize younger children's access to digital media by providing a more natural, voice-based interaction compared to screens and keyboards (Beneteau et al., 2020), it comes as no surprise that already in 2020 the 23% of 2- to 4-year-olds were found to autonomously engage with smart speakers (Rideout & Robb, 2020). One of the first studies observing the use of smart speakers in the domestic context has shown that Amazon Echo was primarily domesticated as a family device and located in a shared space such as the living room (Lopatovska et al., 2019). Moreover, the authors identified three main usage patterns across families—information, entertainment, control of interconnected devices or smart appliances. Consistently, in the Italian survey of families with young children (Zaffaroni et al., 2022), entertainment and information topped the list of activities regularly asked of smart speakers: Respondents reported using Echo or Google Home to listen to music (75%), to get quick factual information (63%), and listen to the news (51%). Smart speakers have also been incorporated into a range of habitual domestic practices, including as a reminder of deadlines and appointments (57%), to access cooking recipes (51%), control connected appliances and energy-saving technologies (51%), communicate with family and friends (43%), and tell bedtime stories to children (43%).

Lopatovska et al.'s (2019) study also pointed to generational differences—with children being more likely to request Echo to play music or tell jokes—and to a decline in both overall usage and specific activities (such as looking for news) over time. That parents and children use smart speakers differently was confirmed by Garg and Sengupta's (2020) study, based on interviews with 18 families and Google Home Activity logs. In fact, while parents used Google Home mainly to play music or control smart lighting, thermostats, and cameras, children's voice logs indicated that playing games, listening to music, searching for information, and engaging in small talk were the most common practices. Analyzing usage patterns over time, the authors observed how parents' scaffolding of children's interaction with the smart speaker was limited in time, until the child learned to adapt their communication style (i.e., by raising their voice, repeating queries, and using shorter sentences). With respect to entity-making practices (Suchman, 2011), Garg and Sengupta (2020) found that, contrary to adults and older siblings, children younger than 7 attributed a human-like

identity to the conversational agent. Over time, this persistent personification resulted in emotional attachment.

Other studies focused on the relationship between different practices of use and the associated family dynamics. Beneteau and colleagues (2020) identified three broad usage patterns that resulted in distinctive parent–child relationships. First, parents and children engaged in practices of co-use, and parents scaffolded their children's interaction with the smart speaker in various ways (i.e., suggesting how to rephrase a question or other repair communication strategies). In this way, younger children's communicative skills were expanded through their parents' scaffolding (what they label as "fostering communication"). Second, conflicts emerged over the control of the smart speaker, with children trying to take control and "disrupt access," whether to impose their desired outputs or simply to annoy their parents. Third, parents engaged in practices of "augmented parenting" (Beneteau et al., 2020, p. 5), using the smart speaker as a neutral third-party mediator in decision-making, or in a more deceitful way, to obtain children's obedience—such as setting an alarm on the smart speaker to enforce bedtime rules. Similar struggles over the control of the smart speakers have been found in a thematic analysis of Chinese user-generated videos (Wang et al., 2023), highlighting how the devices acted as *mediators* or *family members*, contributing to a democratization of child–parent relationships, through an empowerment of the child. Since co-use was found to represent the most popular usage pattern among Chinese families, the authors conclude that smart speakers prompted a shift from "bedroom culture" back to a "living room culture" (Wang et al., 2023, p. 13).

The significance of family's culture in the domestication of smart speakers has been highlighted in a study of the diverse motivations underpinning different usage patterns across different family types (Wald et al., 2023). In fact, different communicative practices were informed by families' distinctive frames of relevance, including trust in technology and preferred parental mediation style, and by their internet literacy.

## Conceptualizing Smart Speakers as Communicative AI

Scholarship in HMC and in mediatization research agree on the properties that distinguish communicative AI from other digital media, and that prompt us to question traditional definitions of both media and (human) communication. In fact, conversational agents, as those embedded in smart speakers, "are not simply media in the sense that they serve as interaction nodes between people" (Hepp, 2020a, p. 79). We now communicate *with* such media, rather than simply *through* them. Indeed, conversational agents, as much as other communicative robots (Hepp, 2020a), are designed to fulfill communicative purposes and simulate human communication capabilities. By expanding the role of digital technology beyond that of a mediator or an interface to that of a communicative partner (Guzman, 2018, 2019; Natale & Guzman, 2022), therefore, conversational agents automate communication.

The current automation of communication is leading to an "increased level of agency in that it is the technology itself that forms and interpret messages" (Natale & Guzman, 2022, p. 630). Not only do conversational agents respond to human inputs, they also collect data on their users and the surrounding environment and distribute them. Thanks to the agency performed through their role of communicators, conversational agents create the

impression of responding and adapting to users' queries in a personalized manner (Natale & Cook, 2020). As a consequence, in automated communicative relationships, human agency is also redefined and constrained: "The agency of the users of voice assistants can be best described as the ability to choose among a pre-defined range of interactions that the companies already anticipated for their systems" (Natale & Cook, 2020, p. 8). However, we should not overestimate the autonomy of automated communicators. Rather, it is more appropriate to think of these agents as partially autonomous automated systems, rather than forms of AI in its narrow sense of sophisticated and complex machine learning systems (Hepp, 2020b). In fact, it's precisely the newly acquired ability to gather, generate, and distribute information on their users that makes automated communication systems depending on the digital-material infrastructures of datafication (Couldry & Mejias, 2019; van Dijck, 2014) for their functioning. On this basis, we should rather speak of quasi-communication (Hepp 2020a, 2020b), to emphasize how these technologies are designed to enter into a communicative relationship with its users, yet are not fully autonomous. Indeed, machine's responses to human inputs are processed and performed by invisible algorithms, software, platforms, and hardware that operate in the background. Therefore, communicative AI is characterized by a double embodiment: "communicative" and "infrastructural" (Hepp, 2020b, p. 1412), with the former operating as to conceal the latter. In fact, the simulation of a human-like conversation helps conceal the collection, transmission, and algorithmic calculation of users' and environmental data on which basis communicative AI operates (Mascheroni & Siibak, 2021). We can thus define the conversational agents embedded in smart speakers, along with Hepp (2020b, p. 1416), as:

> (a) (partially) automated and (partially) autonomous media that (b) serve as quasi-communication interfaces with humans, (c) they are embedded in a comprehensive digital infrastructure and (d) have in most instances an apparently human-like user interface.

Despite their partial autonomy and machine-like nature, the conversational agents integrated in smart speakers are generally perceived and addressed as communicative partners (Guzman, 2018, 2019; Guzman & Lewis, 2020). In fact, conversational agents are programmed to enter the communicative relationship with a specific social positioning, defined by its human-like traits, a gender, and a social role: that of an assistant (Guzman, 2019) or servant (Fortunati et al., 2022). Their anthropomorphization and personification—through the attribution of a name, a naturalistic female voice, and the simulation of an individual personality—is what enables communicative AI to "interact with humans on an emotional level" (Zhao, 2006, p. 408). However, the interpretation of conversational agents as human-like communicators is still ambivalent: Fortunati et al. (2022, p. 8) found that while half of their sample perceived Alexa as an equal or superior communicator, the remaining half understood it as "an inferior communicator (a kind of servant or slave)." Other times, instead, users are more guided by the machine-like nature of conversational agents in their interpretation of their status (Guzman & Lewis, 2020).

These findings suggest that "in any process of communication where humans are involved, humans are ultimately responsible for the construction of sense" (Natale, 2021, p. 908). Second, they confirm that, like any act of communication, automated communication is always socially situated and contingent: While both human and machine agency are

redefined, and the family as a communicative figuration is likely to be transformed in their repertoire of communicative practices and meanings system by the advent of smart speakers (Mascheroni & Siibak, 2021), mediatization is never a linear cause–effect relationship (Hepp, 2020a).

In this article we will address our research questions, firmly grounded in mediatization research (Hepp, 2020b), by building upon the HMC research agenda (Guzman & Lewis, 2020). Accordingly, we will examine the functional and relational dimensions through which families make sense of the conversational agents integrated in smart speakers and relate to these technologies and to themselves. In so doing, we will advance our understanding of how families as communicative figurations change in light of the domestication of smart speakers, and what distinguishes the communicative practices *through* and *with* smart speakers from other mediated communicative practices within the family.

## Methods

### Study Design

This article reports on the findings of three waves of qualitative longitudinal research designed to investigate the datafication of childhood and family life as a socially situated, everyday and embodied experience. Families were recruited through snowball sampling (i.e, digital flyers illustrating the scope of the project were shared on Facebook or WhatsApp local parenting groups, and printed flyers were circulated in different workplaces [hospitals, call centers]). We were interested in families with at least one child aged 8 or younger, because this is the first generation "to be datafied *from birth*" (Children's Commissioner for England, 2018, p. 11). Given the broader focus of the study, owning a smart speaker was not a sampling criterion. Yet, consistently with the survey data collected in September 2020 (Zaffaroni et al., 2022), we found that the majority of families had smart speakers, as illustrated in Table 1. The final theoretical sample consists of 20 families with at least one child aged 8 or younger, living in the city of Milan and its surrounding towns. Only Family 1 dropped out after the first interview. While small in number, and self-selected—their main motivation to take part in the study being receiving guidance or confirmation about children's media use from researchers—the sample represents a diverse demographic range in terms of income and constellation of actors—with many families having multiple children ranging from 0- to 14-year-olds, but also single-parent and/or single-child families (Table 1).

**TABLE 1 Participating Families**

| Family number | Parents (age, nationality) | SES | Selected child on first visit | Siblings | Smart speaker |
|---|---|---|---|---|---|
| Family 1 | Mother (39, Italian)<br>Father (43, Italian) | High | F, 5 | | Amazon Echo |
| Family 2 | Mother (37, Italian)<br>Father (38, Italian) | Medium | M, 4 | M, 1 | Google Home |
| Family 3 | Mother (42, Italian)<br>Father (48, Italian) | Low | F, 3 | | |

| | | | | | |
|---|---|---|---|---|---|
| Family 4 | Mother (38, Russian)<br>Father (38, Italian) | Medium | F, 4 | | Amazon Echo + Alexa (Smart TV) |
| Family 5 | Mother (37, Belgian)<br>Father (45, Italian) | High | F, 6 | M, 3 | |
| Family 6 | Mother (43, Italian)<br>Father (43, Italian) | Medium | M, 5 | | Google Home<br>Amazon Echo (third wave) |
| Family 7 | Mother (42, Italian-Swiss)<br>Father (39, Italian-French) | Medium/ High | M, 5 | F, 2 | |
| Family 8 | Mother (41, Italian)<br>Separated | Medium | M, 5 | | Amazon Echo (dropped out) |
| Family 9 | Mother (41, Italian)<br>Father (42, Italian) | High | M, 7 | M, 18m; M, 3; M, 6; F, 10; F, 13; F, 14; | |
| Family 10 | Mother (40, Italian)<br>Father (44, Italian) | Medium | M, 7 | | Amazon Echo (dropped out) |
| Family 11 | Mother (34, Moroccan)<br>Father (46, Italian) | Low | M, 6 | M, 8 | |
| Family 12 | Mother (38, Italian)<br>Separated | Medium | M, 6 | F, 10 | |
| Family 13 | Mother (41, Italian)<br>Father (49, Italian) | Low | M, 6 | | Amazon Echo |
| Family 14 | Mother (40, Italian)<br>Father (40, Italian) | Medium | F, 7 | M, 3; M, 12; F, 10 | Google Home |
| Family 15 | Mother (42, Italian)<br>Father (42, Italian) | Medium/ Low | M, 6 | | |
| Family 16 | Mother (40, Italian)<br>Divorced | Medium | M, 7 | M, 5 | Amazon Echo (at father's home) |
| Family 17 | Mother (37, Moldavian-Russian)<br>Separated | Low | F, 8 | | Amazon Echo (at father's home) |
| Family 18 | Mother (40, Italian)<br>Father (41, Italian) | Medium/ High | F, 8 | F, 10 | Amazon Echo |
| Family 19 | Mother (53, Italian)<br>Father, (58, Italian) | High | M, 5 | | Amazon Echo |
| Family 20 | Mother (49, Italian)<br>Father (49, Italian) | Medium | F, 8 | F, 11 | Amazon Echo |

Overall, 58 interviews were conducted face-to-face, at home, during three waves of data collection (November–December 2021, April–May 2022, and December 2022–January 2023). During the first family visit, parents and children were initially involved in an ice-breaking activity; next, one researcher followed the child in a toy and digital media tour (Plowman & Stevenson, 2013), while the second interviewed parents about the family's media practices, parental mediation, their expectations and fears around digital media and IoTs. In the second wave, the researchers conducted group interviews at home, using the map drawing method (Watson et al., 2022) as a visual prompt to stimulate discussion about family routines, the domestication of digital media, power relations around media and material space. At the end of the second interview, parents were invited to fill in a digital media diary for 1 week, recording the selected child's interactions with digital media each day (Mascheroni & Zaffaroni, 2023). On the third visit, researchers and interviewees engaged reflexively on the findings of the first and second wave, visualized in a map through a network analysis of actors, media, and interactions (Amadori, 2023). The research protocol was approved by the University's IRB. Parents signed an informed consent, and researchers paid attention to children's assent—understood as a process of ongoing negotiation between the researcher and the child-participant (Warin, 2011). In this article, we will focus on interview data from families with smart speakers ($N = 37$).

### Data Analysis

Interviews were transcribed, anonymized, and analyzed using MaxQDA2022. The pseudonyms used in the transcriptions were chosen by the participating children. In line with constructivist grounded theory (CGT) (Charmaz, 2014), the data analysis process was iterative. First, the same interview transcript was coded independently by each researcher, using line by line inductive coding. The comparison between the coded transcripts allowed to develop a shared code sheet, which informed the second stage of analysis, when the interviews were coded independently by researchers. The process was iterative and the shared code sheet was enriched and revised until each coder was satisfied. Finally, the PI examined all coded transcripts and finalized a second-level coding by revising codes and aggregating them into broader codes, summarizing the main themes emerged from interviews and informed by the theoretical framework. The coding scheme was further refined and enriched during the analysis of the second and third waves of interviews.

## Findings

### Incorporating Smart Speakers in the Family's Media Ensemble and Repertoire of Practices

Nine families in our sample own and use a smart speaker, located in the kitchen or the living room. Two families dropped out for privacy concerns (Family 10), or concerns over the child-smart speaker interaction (Family 8) respectively. In two divorced families, children have access to a smart speaker only in the father's home (Table 1).

As a new addition to each family's media ensemble, the smart speaker may entertain relationships of complementarity, substitution, or control with other media, as well as variously fit the family's communication repertoire. In our participating families, the use of smart speakers is mainly incorporated into the pre-existing communication and media repertoires, with the control function representing the only new practice. In fact, the most frequent usage practices fit into the broad use genres of entertainment (music, fairy tales, jokes), information (quick factual information, news, weather forecast, schoolwork, cooking recipes), and control (of other media or appliances). Some generational and gender differences emerge (as we will further elaborate below in relation to power dynamics between family members): children are more likely to engage with the smart speaker to listen to music, ask for fairy tales or jokes, and practice small talk, while adults ask for information and listen to music; fathers are responsible for the control functions.

Moreover, we mainly found relations of complementarity, without significant disruptions to the pre-existing media ensemble. For example, children's requests to smart speakers, such as narrated stories or music, do not necessarily replace pre-existing media practices (e.g., books or YouTube), but rather complement them as a convenient alternative due to the voice activation. Only when located in the kitchen (Families 2 and 14), and limited to adults' practices, smart speakers replace radios or smartphones for news, cooking timer, and recipes. The relationship of control, in its strict sense, is limited: In the third wave only Family 18 continued operating smart lights via a smart speaker, whereas Family 14 dropped out of lights automation—due to the mother repeatedly clicking the switch—while Family 13 occasionally used the remote activation of smart lighting when on holiday—to prevent burglaries.

Importantly, however, while smart speakers are incorporated within pre-existing mediatized family practices, as interfaces to access and control other media and platforms (e.g., Spotify or the morning news on radio stations), they are implicated in a complex network of interconnected devices and services, which collect, track, and exchange data. Therefore, while smart speakers broaden the range of media through which families carry out their usual media or communication practices, due to their infrastructural embodiment they simultaneously intensify the family's reliance on data relations for the performance of many mundane (individual or shared) activities.

## Changing media ensemble, changing family figurations?

Smart speakers set new conditions for agency within the family, subverting or reinforcing existing power relations in the home through the communicative practices they enable. In terms of intergenerational power relations, parents are challenged in their parental mediation role, as gatekeepers of children's access and use of digital media. In fact, even pre-schoolers are now able to access media content autonomously, through voice commands, thus gaining control not only on the smart speaker but also on the interconnected devices (e.g., the smart TV). The novel autonomous engagement with media achieved by young children—and the seemingly democratized child–parent relations (see Wang et al., 2023)—is not necessarily restricted nor disapproved by parents. For example, the mother in Family 2 is proud of her son's greater acquired agency, and encourages him to interact with the smart speaker:

> Petra [mother, Family 2, wave 1]: It's nice because he [4-year-old son] asks it for songs. I mean, "Ok Google, I want to listen to the soundtrack of Moana," I don't know . . . or "I want to listen to . . .". It's nice because he can ask for the songs on his own, and he does. Also, every morning he asks Google what the weather will be like in Milan. I used to do it before, but now he does it. [. . .] sometimes it is me telling him "let's ask Google what's the weather like today."

The communicative interaction described in the excerpt above, that has become an integral part of the morning wake-up routine, shows how the smart speaker involved a shift in power, from the mother to the child, in the relationship with information: now it is the child who is responsible for accessing the weather forecasts, although often prompted by his mother.

Moreover, although smart speakers may actually enable children to bypass parental mediation, the empowerment of the child does not necessarily translate into a corresponding disempowerment of parents. Power is not always a zero-sum game within family figurations. Children's voice-based autonomy, in fact, is intertwined with and dependent on how agential practices unfold on the side of parents. In certain family configurations, the child's empowerment is paralleled by a corresponding benefit for parents, who equally turn the smart speaker into a technology of empowerment. This can be observed, for example, when the smart speaker serves as a digital babysitter, to keep the child occupied with stories and music. Petra [mother, Family 2, wave 1] further explains that she likes her son to be able to ask Google Home autonomously for stories because "Spotify is full of nice things . . . and when I need to relax, or I need him to be still, I put on a lot of audio stories."

Likewise, by granting autonomous access to music through voice commands, smart speakers alleviate single parents from the preoccupation of constantly supervising their children's screen time. Letizia recounts having bought Amazon Echo a few months before our third visit to avoid her 5-year-old son accessing her smartphone whenever he wanted to listen to music:

> Letizia [mother, Family 8, wave 3]: I made the decision to buy it unconcernedly because I thought could be, even for him, something convenient, because he did no longer have to access my phone . . .

Often, however, the domestication of smart speakers poses a direct challenge to parents. In fact, by enabling a more autonomous access to media content, smart speaker risk undermines parents' agency in preventing children's exposure to inappropriate content. For example, Family 1 explains how they resorted to the employment of technical mediation (namely, parental controls), to protect their daughter from inappropriate content (e.g., *fart jokes* that Alexa prompted to a 5-year-old girl):

> Claudia [mother, Family 1]: The other day, when you were not at home, she [daughter] was asking songs to Alexa, and at some point I heard Alexa replying "I do not have that title, but try with 'fart something,' but you need to unblock it on the app first."

While additional parental mediation helps counterbalance the quasi-autonomous agency of the smart speaker, that suggests content directly challenging the family's frame of relevance, other times it is children's smart speaker-enabled empowerment that can lead to overt conflicts between family members. This happens when transformations in pre-existing power relationships and communicative practices (*with* and *through* media) clash with the family's shared norms and reciprocal expectations—namely, when the communicative practices enabled by smart speakers clash with the family's frame of relevance. For example, Camilla [mother, Family 20, wave 3] recounts when, during a play date, a friend of her son, whose mother is "a zealot" of age-appropriate content, asked the smart speaker to play the Måneskin [Italian rock band] and "danced shirtless! It made me laugh!" Gabriella, instead [mother, Family 6, wave 2], was annoyed when "[her 5-year-old son] interrupted Google without asking me, because he wanted his songs." Arguments of this kind continued over time: "He still does it. Especially if he wants to watch TV and I am listening to music, he turns the smart speaker off and watches TV. So 'Alexa, stop!' And so we have an argument because I'd like to listen to music. Or [we argue] about the kind of music." [Gabriella, Family 6, wave 3]. Similarly, Letizia recalls how the daily conflicts over the control of the smart speaker were one of the reasons that motivated her decision to hide Echo in a drawer:

> Letizia [mother, Family 8, wave 3]: I was no longer free to listen to music because he would get there and change it [the child scoffing in the background]. He wanted to change song. It was a continuous war. So I told myself "enough!" Sometimes you listen to your music, sometimes I listen to mine, but we cannot change it continuously. It became unnerving.

Such conflicts over the smart speaker suggest that the point of "disrupting access" (Beneteau et al., 2020, p. 5) may be less about reclaiming identity through personal music choices, and more about resisting parental mediation and reaffirming more balanced power relations in the parent–child dyad. When we look at the gender relationships in the family, instead, we can observe how smart speakers tend to reproduce the gendered power differentials in the home—with fathers being in charge of the management of accounts (as in the example of Family 1 above), as well as of the control functions of smart speakers over smart home solutions (e.g., lighting), and the mothers being in a subordinate position of passive or unskilled users:

> Sara [mother, Family 18, wave 1]: Don't know, my sister uses it for everything, to remind her to hang the laundry or turn off the . . .
>
> Beatrice [8-year-old daughter]: She even turns the TV on!
>
> Sara: While it does not usually come to my mind the idea of using the smart speaker for . . . Just the lights, that then, if I accidentally use the switches, Alexa does no longer listen to voice commands, and my husband gets mad at me!

In certain respects, then, the negotiations and conflicts around smart speakers are not dissimilar to those occurring whenever a new medium enters the family's media ensemble: in fact, the latest addition to the media ensemble is appropriated within the family's

communicative repertoire through the lens of the family's relevance frame, resulting in a, at least partial, reconfiguration of the family as a communicative figuration. What is novel, therefore, is not the ways in which power relations are reproduced, reconfigured, or resisted through smart speakers-supported practices per se. Rather, it is the communication with smart speakers, their being media *through* and *with* which the family communicates, that generates new power dynamics as we will elaborate in the next section.

However, the automated, data-based nature of smart speakers renders their domestication—and the consequent adaptation of the family figuration to their adoption—somewhat distinctive. Indeed, alongside their "communicative" embodiment (Hepp, 2020b), smart speakers are embodied in the infrastructures of datafication. In fact, underlying the functioning of smart speakers is the real time algorithmic processing of users' voice inputs, their classification and matching with the desired output. The incorporation of smart speakers into the domestic context thus entails an intensified mediatization of everyday life, in the form of its progressive datafication and algorithmization. This yields important shifts in the power dynamics between humans and machines, and, consequently, in the family as a communicative figuration.

As already described, by introducing a further level of mediatization in the family's media and communication repertoire (e.g., access to music is now mediated by a voice interface, that activates the playlist on Spotify) smart speakers gain greater agency, that results in a corresponding diminished agency for users. In fact, instead of browsing from a potentially infinite library of contents, users are directly presented with content, and left with little control over the choice. By repeating the outcomes of previous interactions and narrowing down the range of content, smart speakers hinder user's ability to make autonomous choices:

> Tommaso [4-year-old son, Family 2, wave 2]: Ok Google, I want to listen to children's stories!
>
> Google: I play children's stories from Spotify.
>
> Petra [mother, Family 2]: Let's see which one Google chooses today. . . . [Google: "Rapunzel . . ."] that's it! Since it's one we've often asked, because Tommaso likes it, [Google] often plays Rapunzel when it has to choose.

This algorithmization of everyday practices lies at the heart of the tension between the empowerment and disempowerment of users, and the agency of machines. Smart speakers acquire a crucial role with respect to the distribution of agency within the family thanks to the datafication and subsequent algorithmic calculation of everyday life. Sometimes, users feel restricted in their power to choose media content, as Pamela, who explaining why their use of Echo diminished over time, tells:

> Pamela [mother, Family 13, wave 1]: It is useless. Music, anyway, is not unlimited . . . Because it is connected to Amazon, so the library is limited. Instead, I know that Google Nest is connected to YouTube, so there's everything. Conversely, Echo is very limited. Furthermore, it does not understand many things. After all, she is not human!"

Contrary to Pamela, however, most children and parents are unaware of how this processing works and are confronted with content over which they have no direct control. Pamela's excerpt above is interesting also because of her suggestion that smart speakers' limitations depend on its machine-like nature. This comment leads us to the second research question; that is, how communication through and with smart speakers is different from other mediated communication.

## Communicating *With* (and *Through)* Smart Speakers

Besides shifting the power relations between family members, and between the constellation of actors and the media ensemble, smart speakers also afford the, at least partial, construction of a communicative relationship. Indeed, although conversations with smart speakers are brief—and tend to decline over time, when the "novelty effect" fades—new communicative practices emerge, with family members now interacting with a gendered voice trained to perform various tasks, as well as respond to curiosities and jokes.

The communicative interaction with smart speakers is shaped, first, by the attribution of a human-like or, alternatively, a thing-like identity to the technology; and, second, by how effectively the smart speaker responds to the user's request. The two aspects are in fact interrelated: indeed, communicative failures are interpreted through the lens of the identity attributed to the conversational agent integrated in the smart speaker. As a consequence of the mutual shaping of users' perceptions and technology's responses, communication with smart speakers is always socially situated, contingent upon entity-making practices (Suchman, 2011) and the specific interactive occurrence.

Younger children tend to personify voice-based assistants and engage more frequently in entity-making practices, testing the conversational agent's "liveness." Our interviews and observations include numerous examples of children enlivening the smart speaker through their interactions, by constructing the smart speaker as a living entity and attributing it human-like activities and emotions. For example, children asked Alexa if she liked Nutella—as Elisa (8-year-old, Family 20, wave 3)—or if she had parents, as in the following excerpt:

> Camilla [mother, Family 19, wave 2]: One day I was in the kitchen and I heard him talking to someone, and I said, "Who is he talking to?" He was talking to, he was asking Alexa, "Alexa, do you have friends? Who are your daddy and your mommy?"
>
> Alessandro [5-year-old son]: "Alexa! Do you have a mom or a dad?" [repeating Alexa's words] "She was invented . . ."
>
> Interviewer: ". . . by a team of engineers at Amazon."
>
> Alessandro: Yes, it doesn't have parents, it's normal!

The personification of smart speakers through an attribution of anthropomorphic qualities and a personality is largely dependent on its built-in gendered traits. Children, and parents alike, attributed a different gender identity to the conversational agents embedded in Echo or Google; interestingly, this gender difference leads to conflicting and stereotyped

interpretations of their distinctive qualities. For example, the four siblings in Family 14 argued whether Alexa or Google was better. In the conversation, Alexa was attributed greater emotional and conversational qualities, whereas Google was perceived as more intelligent. Only the older brother, who attributes a machine-like quality to the smart speaker, thinks that they perform the same basic functions:

> Alice [10-year-old sister, Family 14, wave 2]: I would prefer Alexa, because she is much more responsive.
>
> Martino [12-year-old brother]: It's the same!
>
> Alice: No, because if you tell Alexa "do you know that I love you so much?" she goes "Thank you for your loooove" [singing].[. . .]
>
> Interviewer: So, why did your dad chose Google Home over Amazon Echo?
>
> Carlotta [7-year-old sister]: Because Google is a male, and knows more things.

Importantly, as anticipated, the attribution of anthropomorphic and gendered characteristics forms part of the communicative labor that families perform to understand and interpret the smart speakers' communicative failures. For example, when complaining about the failures of the smart speaker to locate content across different SVOD platforms, Ludmilla (mother, Family 4, wave 2) mobilized the female servant stereotype: "ours is Alexa's maid, because she doesn't understand!"

Similarly, when Family 6 replaced Google Home with an Echo before our third visit—with Google being downgraded to the bathroom—the different perceived performances of each were interpreted in terms of gender differentials, with women being acknowledged as smarter than men:

> Gabriella [mother, Family 6, wave 3]: By the way, we now have Alexa. We bought Alexa to replace Google, so we eventually have a woman instead of a man and she does understand much more! [laughing]

In Family 16, Amazon Echo was attributed the status of a female competitor who always took the side of her husband. In fact, Alice (mother, Family 16, wave 1) recounts how she "did not get along with Alexa, she did not understand me. Don't know, she probably hated me!" However, despite the fact that parents themselves mobilize anthropomorphic frames in trying to make sense of the occasional communicative failures, the entity making practices in which children engage can also be a source of concern for parents. Letizia explains she stopped using Echo after she observed her son attributing the smart speaker a human-like, yet inferior, status. First, the young child's personification of the smart speaker generated feelings of discomfort and anxiety:

> Letizia [mother, Family 8, wave 3]: I took it away because it was impossible to keep it in the house with him. I didn't like it because of the orders. No, no. You order Alexa, but I just didn't like [his] way of approaching it. [. . .]

> Lodovico [5-year-old son, Family 8]: I did not like it either, because sometimes I asked for a song and it played another one. It did not understand anything!
>
> Interviewer: And this made you angry?
>
> Lodovico: Yes, I wanted to flush it through the toilet. She is stupid.
>
> Letizia: Mmm no, you have to understand it is a machine, not a person. While I observed him . . . trust me, I was scared. It is too soon. [. . .] It upset me. It is normal and automatic for us. But I have noticed it was different for him, he conversed with the smart speaker as if . . . another person was at home. I don't know, I felt uncomfortable about it.

Second, Lodovico extended the inferior status attributed to the smart speaker also to other family members, thus threatening the family's values and norms. Therefore, besides conflicts over the control of the interface, the smart speaker was de-domesticated because of the child's habit to give orders to his mother overtly conflicting with the family's frames of relevance.

As these examples illustrate, constructing smart speakers as human-like communicative partners yields partial or unsatisfactory outcomes, as both children's and parents' expectations of reciprocity and mutual understanding are often frustrated by the limited capacities of the device, due to its faulty algorithmic processing of voice inputs—failures that may be due to lower training of voice recognition in languages other than English. After the novelty effect fades, children also grow tired of the company of smart speakers, to the point that they forget their presence:

> Pamela [mother, Family 13, wave 1]: At the beginning he [6-year-old son] was [attracted], he would ask Alexa for Google searches, "Alexa what is the biggest animal in the world?" Sometimes we still do this, just for fun, "Alexa, what is the most stinky animal in the world?" You know, stupid things . . . but very seldom.

Conversely, we observed only one example of communication *through* smart speakers, when Family 20 tested the possibility of voice calls via the smart speaker "for security reasons," to have an additional communication channel "in case mobile phones did not work" [Umberto, father, Family 20, wave 3].

## Conclusions

This study aimed to understand whether and how families, as communicative figurations, change through the domestication of smart speakers, and what is distinctive about the communicative practices *through* and *with* smart speakers from other mediated communicative practices within the mediatized home. Our findings show that, while the domestication of smart speakers follows similar patterns across families—with their incorporation into the pre-existing media repertoire as a further digital interface to habitually consumed content, followed by a general decline in usage over time—interactions with smart speakers are also distinctive of each family configuration and the type of device, depending on the situated

nature of their domestication. More specifically, the family's communication repertoire is enhanced by the relationships of complementarity and control (Wang et al., 2023) that the smart speaker entertains with the pre-existing media ensemble. Substitution is less frequently observed and limited to the smart speaker functioning as a new interface to routine media content such as music (playlists) and news.

In terms of reconfigurations of the relationships between family members, smart speakers are technologies of both empowerment and disempowerment. Children are rendered more autonomous in their media choices, thus challenging parental mediation and reclaiming agency through acts of "disrupting access" (Beneteau et al., 2020, p. 5). However, the empowerment of the child does not necessarily translate into a corresponding disempowerment of parents, who could benefit from smart speakers as digital babysitters, or from voice interfaces as an alternative to screens. If intergenerational relationships are re-negotiated, often gendered power imbalances are instead reproduced, with fathers exercising control over the smart speaker and interconnected devices.

Overall, our findings emphasise how smart speakers continue and intensify the mediatization of families' everyday lives (Mascheroni & Siibak, 2021). While power negotiations and redistribution of agency within families have always been characteristic of the domestication of media (Silverstone, 1999), what is distinctive about smart speakers is their double embodiment (communicative and infrastructural) (Hepp, 2020b), and the relationship between the two. The communicative embodiment of smart speakers is crucial in their domestication: in fact, although children are more likely to engage in entity-making practices (Suchman, 2011), the attribution of anthropomorphic and gendered characteristics forms part of the communicative labor that all family members perform to make sense of the smart speakers, including its failures. The interpretations of both the human-like and the machine-like attributes of smart speakers changes according to the technology's ability to complete a task and understand human inputs. Consequently, the smart speaker's positioning as a communicative partner is dynamic and contingent upon the specific interactional situation, shifting along the continuum of *human*-like and *thing*-like based on how human expectations are satisfied or frustrated. Rather than new family members (Wang et al., 2023), across our families smart speakers are either perceived as media or, less often and primarily by younger children, as companions, equal to pets or toys.

Smart speakers' infrastructural embodiment also emerged as a key aspect in the domestication practices of our participating families. As interconnected devices able to control other media and appliances, smart speakers acquire an agentic role by further extending the datafication and algorithmic calculation of everyday life. This yields important shifts in the power dynamics between humans and machines, with humans having little control over the choice of media content and, more generally, over the extraction, distribution, and analysis of users' and home data. Indeed, the human-like communicative embodiment materialized in a gendered voice conceals the complementary embodiment of the device in the digital-material infrastructures of datafication. More than sustaining new power relations among the actors' constellation distinctive of each family configuration, smart speakers involve a new relationship between human actors and artifacts, that Couldry and Hepp (2017, p. 131–132 ) called "tool reversability." Tool reversibility indicates how users themselves are being used by the data-based, internet-connected tools embedded in their

everyday lives. Therefore, while smart speakers are appropriated as technologies of empowerment, they equally remain technologies of power, depending on our agency and simultaneously subtracting part of our agency.

The current study has a number of limitations, some of which are common to qualitative research (a diverse but not representative sample of parents who volunteered to take part in the study for their own interest in and concerns around children's digital media use), and others specific of our research design. In fact, our research project was not focused on the domestication of smart speakers, but on mediatized and datafied families as communicative figurations. Additionally, except for observing a general decline in usage, processes of de-domestication (Family 8), or the domestication of a new smart speaker combined with the re-domestication of the older device (Family 6), the longitudinal nature of our study was not further explored. Future studies could address this limitation and investigate whether the shifting power dynamics here described are distinctive of families with young children, or whether a more substantial re- or de-domestication occurs over a longer time frame.

## Author Biography

**Giovanna Mascheroni,** (PhD, Università Cattolica del Sacro Cuore) is a sociologist of digital media, and Professor in the Department of Communication, Università Cattolica del Sacro Cuore. She was the leader of WP6 in the H2020 project ySKILLS, and the PI of DataChildFutures, a national project investigating the data practices of Italian families with children aged 0- to 8-year-olds. Her work focuses on the social consequences of digital media, datafication, and AI for children and young people. Her latest book is *Datafied childhoods: Data practices and imaginaries in children's lives* (2021), co-authored with Andra Siibak.

https://orcid.org/0000-0002-6939-2650

## References

Amadori, G. (2023). Studying the datafied home: The potentials of combining mixed methods, constructivist grounded theory and social network analysis. *ZeMKI Communicative Figurations Working Paper 44*. https://web.archive.org/web/20240102154030/https://www.uni-bremen.de/en/zemki/publications/working-paper-series/

Beneteau, E., Boone, A., Wu, Y., Kientz, J., Yip, J., & Hiniker, A. (2020). Parenting with Alexa: Exploring the introduction of smart speakers on family dynamics. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, April 21, 1–13. https://doi.org/10.1145/3313831.3376344

Charmaz, K. (2014). *Constructing grounded theory: A practical guide through qualitative analysis. Second Edition*. Sage.

Children's Commissioner for England. (2018). *Who knows what about me? A children's commissioner report into the collection and sharing of children's data*. https://web.archive.org/web/20190212122718/www.childrenscommissioner.gov.uk/wp-content/uploads/2018/11/who-knows-what-about-me.pdf

Couldry, N., & Hepp, A. (2017). *The mediated construction of reality*. Polity.

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Fortunati, L., Edwards, A., Edwards, C., Manganelli, A. M., & de Luca, F. (2022). Is Alexa female, male, or neutral? A cross-national and cross-gender comparison of perceptions of Alexa's gender and status as a communicator. *Computers in Human Behavior, 137*, 107426. https://doi.org/10.1016/j.chb.2022.107426

Garg, R., & Sengupta, S. (2020). He is just like me: A study of the long-term use of smart speakers by parents and children. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4*(1), 1–24. https://doi.org/10.1145/3381002

Guzman, A. L. (2018). What is human–machine communication, anyway? In L. Guzman (Ed.), *Human–machine communication: Rethinking communication, technology, and ourselves* (pp. 1–28). Peter Lang.

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior, 90*, 343–350. https://doi.org/10.1016/j.chb.2008.03.008

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human-Machine Communication research agenda. *New Media & Society, 22*(1), 70–86. https://doi.org/10.1177/1461444819858691

Hepp, A. (2020a). *Deep mediatization*. Routledge.

Hepp, A. (2020b). Artificial companions, social bots and work bots: Communicative robots as research objects of media and communication studies. *Media, Culture & Society, 42*(7–8), 1410–1426. https://doi.org/10.1177/0163443720916412

Hepp, A., & Hasebrink, U. (2018). Researching transforming communications in times of deep mediatization: A figurational approach. In A. Hepp, A. Breiter, & U. Hasebrink (Eds.), *Communicative figurations: Transforming communications in times of deep mediatization* (pp. 15–48). Springer. https://doi.org/10.1007/978-3-319-65584-0_2

Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., William, H., Sorsche, P., Hirsch, D., Li, Q., & Martinez, A. (2019). Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science, 51*(4), 984–997. https://doi.org/10.1177/0961000618759414

Mascheroni, G., & Siibak, A. (2021). *Datafied childhoods. Data practices and imaginaries in children's lives*. Peter Lang.

Mascheroni, G., & Zaffaroni, L. G. (2023). From "screen time" to screen times: Measuring the temporality of media use in the messy reality of family life. *Communications*. https://doi.org/10.1515/commun-2022-0097

Natale, S. (2021). Communicating through or communicating with: Approaching artificial intelligence from a communication and media studies perspective. *Communication Theory*, *31*(4), 905–910. https://doi.org/10.1093/ct/qtaa022

Natale, S., & Cook, H. (2020). Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture & Society*. https://doi.org/10.1177/0163443720983295

Natale, S., & Guzman, A. L. (2022). Reclaiming the human in machine cultures: Introduction.

*Media, Culture & Society, 44*(4), 627–637. https://doi.org/10.1177/01634437221099614

Plowman, L., & Stevenson, O. (2013). Exploring the quotidian in young children's lives at home. *Home Cultures, 10*(3), 329–348. https://doi.org/10.2752/175174213X13739735973381

Pridmore, J., Zimmer, M., Vitak, J., Mols, A., Trottier, D., Kumar, P. C., & Liao, Y. (2019). Intelligent personal assistants and the intercultural negotiations of dataveillance in platformed households. *Surveillance & Society, 17*(1/2), 125–131. https://doi.org/10.24908/ss.v17i1/2.12936

Rideout, V., & Robb, M. B. (2020). *The common sense census: Media use by kids age zero to eight, 2020*. Common Sense Media. https://web.archive.org/web/20201117140451/www.commonsensemedia.org/research/the-common-sense-census-media-use-by-kids-age-zero-to-eight-2020

Silverstone, R. (1999). *Why study the media?* Sage.

Suchman, L. (2011). Subject objects. *Feminist theory, 12*(2), 119–145. https://doi.org/10.1177/1464700111404205

van Dijck, J. (2014). Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance and Society, 12*(2), 197–208. https://doi.org/10.24908/ss.v12i2.4776

Wald, R., Piotrowski, J. T., Araujo, T., & van Oosten, J. M. (2023). Virtual assistants in the family home. Understanding parents' motivations to use virtual assistants with their child(dren). *Computers in Human Behavior, 139*, 107526. https://doi.org/10.1016/j.chb.2022.107526

Wang, B., Luo, L., & Wang, X. (2023). "Back to the living room era": Smart speaker usage and family democracy from the family dynamic perspective. *New Media & Society*. https://doi.org/10.1177/14614448231155624

Warin, J. (2011). Ethical mindfulness and reflexivity. *Qualitative Inquiry, 17*(9), 805–814. https://doi.org/10.1177/1077800411423196

Watson, A., Lupton, D., & Michael, M. (2022). The presence and perceptibility of personal digital data: Findings from a participant map drawing method. *Visual Studies*. https://doi.org/10.1080/1472586X.2022.2043774

Williams, R. (2003 [1974]). *Television: Technology and cultural form*. Routledge.

Zaffaroni, L. G., Amadori, G., & Mascheroni, G. (2022). *DataChildFutures. Rapporto sui risultati dell'indagine 2020*. https://doi.org/10.5281/zenodo.6367526

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society, 8*(3), 401–419. https://doi.org/10.1177/1461444806061951

HUMAN-MACHINE
COMMUNICATION