UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
The effect of gender on multimodal child-directed language: Evidence from analyses of broadcast programmes

Permalink
https://escholarship.org/uc/item/86v6q19b

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors
Zhang, Yanran
Han, Mengru
Swerts, Marc
et al.

Publication Date
2025

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

The effect of gender on multimodal child-directed language: Evidence from analyses of broadcast programmes

Yanran Zhang (Y.zhang@tilburguniversity.edu)
Department of Communication and Cognition, Tilburg University, Warandelaan 2, Tilburg, 5037 AB Netherlands

Mengru Han (mrhan@zhwx.ecnu.edu.cn)
Department of Chinese Language and Literature, East China Normal University,500 Dongchuan Rd, Shanghai, 200241 China

Marc Swerts (m.g.j.swerts@tilburguniversity.edu)
Department of Communication and Cognition, Tilburg University, Warandelaan 2, Tilburg, 5037 AB Netherlands

Yan Gu (yan.gu@essex.ac.uk)
Department of Psychology, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ UK
Experimental Psychology, University College London, 26 Bedford Way, London, WC1H 0AP UK

Abstract

We investigated gender differences in multimodal communication directed to children and adults. Eighty-two broadcasters (46 females and 36 males) participated in hosting adult-directed and child-directed broadcasting programmes respectively, and their lexical/syntactic features, prosody, and gestures were compared. Results revealed that broadcasters adapted their communication styles when addressing children. However, notable gender differences emerged: male broadcasters exhibited less diverse vocabulary, longer utterances, lower pitch but higher intensity, faster speaking rate with more pauses, and fewer referential gestures than their female counterparts. Furthermore, male broadcasters demonstrated larger adjustments in word frequency and vocal intensity but smaller adjustments in the use of questions and gestures than females. These findings highlight distinct patterns in how men and women adapt multimodal communication to children, offering valuable insights into gendered strategies in child-directed language production and recipient design. Moreover, it offers implications for developing tailored broadcast training.
Keywords: child-directed language; gender differences; multimodal language; broadcasting; gesture; prosody

Introduction

Children are exposed to both child-directed language (CDL) and adult-directed language (ADL) in their daily lives. CDL is characterized by specific multimodal features, such as more salient prosody (e.g., Han, de Jong, & Kager, 2022), simplified vocabulary, many repetitions and shorter utterances, and more frequent use of representational gestures than ADL (e.g., Campis & Özyürek, 2013; Crista, 2013; Fernald, 1992; Gallaway & Richards, 1994; Snow, 1995; Soderstrom, 2007). These features have been shown to facilitate children's language learning (e.g., Dong, Gu, & Vigliocco, 2021; Donnellan et al., 2023; Shi, Gu, & Vigliocco, 2023; Han, de Jong, & Kager, 2023). Notably, CDL appears to be a common phenomenon across cultures and languages (Cox et al., 2022), being used by parents, caregivers and adults without children (e.g., Cebioglu, Marin, & Broesch, 2021; Jacobson et al., 1983; Rowe, 2012; Shneidman & Goldin-Meadow, 2012).

While CDL is considered universal, the degree of adjustments varies across individuals. Research indicates that CDL is influenced by factors such as family socioeconomic status (Romeo et al., 2018), education levels (Rowe, 2008), personality traits (Zhang & Gu, 2024) and parental sensitivity (Leerkes et al., 2009; Spinelli & Mesman, 2018). Moreover, gender-related factors may contribute to variation in CDL. Most studies on CDL have focused on mothers or female caregivers due to their traditional central role in parenting. With the increasing involvement of fathers and male caregivers, recent research has started examining gender differences in CDL behaviours, yielding mixed findings.

A few studies suggest that men generally talk less frequently and for shorter durations than women (Davidson & Snow, 1996) and that mothers initiate more conversations with children at home than fathers (VanDam et al., 2022). Additionally, female caregivers tend to maintain closer proximity to children compared to male caregivers (Kristensen et al., 2020). Lexical and syntactic features of CDL also show gender-based differences. For instance, women's CDL has been found to be more elaborate, with longer mean utterance lengths (MLU), greater lexical diversity, longer turn lengths, and more frequent use of Wh-questions than men's (Majorano et al., 2013; Malone & Guy, 1982; Pancsofar & Vernon-Feagans, 2006; Rondal, 1980). However, other studies reported no significant differences in lexical diversity or MLU between mothers and fathers after accounting for the amount of parental language output or in low-income families (Pancsofar & Vernon-Feagans, 2006; Rowe et al., 2004) or between males and females in same-sex-parent families (Grinberg et al., 2022). Prosodic adjustments also vary: fathers tend to modulate prosody less than mothers when interacting with toddlers (Kondaurova et al., 2023), though socially less competent males are more likely to adjust their child-directed prosody (Kempe, 2009), as the trade-off between mating effort and parental effort showed that less attractive men invest more in their children than more attractive men (Gangstead & Simpson, 2000). Additionally, contrary findings showed that both fathers and mothers had exaggerated prosody and expanded vowel

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

spaces (Rosslund et al., 2024), and fathers even increased their pitch variability within and across utterances more than mothers did (Benders, StGeorge, & Fletcher, 2021).

While lexical, syntactic and prosodic features of CDL have been studied extensively, less attention has been given to non-verbal behaviours, such as gestures, including pointing (indicating a referent), representational gestures (iconic depiction), and beat gestures (aligned with speech rhythm) (McNeill, 1992). Gestures also exhibit notable gender differences. Females use more diverse (different types) and frequent gestures than males (e.g., Kavakli & Chen, 2014; Nicoladis et al., 2007). This raises the question of whether such gender differences in gestures also influence the way people adapt their communication in CDL, potentially impacting children's language development.

Studying CDL multimodally is crucial for several reasons. First, language is inherently multimodal, with words, prosody, and gestures all contributing to communication (e.g., Hagoort & Özyürek, 2024; Motamedi et al., 2024). Research on CDL has shown that non-verbal behaviours such as pointing, beats, gazing and touching facilitate children's word learning and memory (Abu-zhaya, Seidl, & Cristia, 2017; Cameron & Xu, 2011; McGregor, 2008; So et al., 2012). A multimodal approach thus provides a more comprehensive understanding of gender differences in language input and their possible impact on language comprehension and learning. Second, multimodal cues enhance engagement, which plays a key role in child interaction. For example, infants find onomatopoeia and sound effects amusing (Dingemanse & Thompson, 2020), respond more positively to female smiley faces (Bayet et al., 2015), and show better joint attention when exposed to gazing and pointing (Brooks & Meltzoff, 2008; Vigliocco et al., 2019) or prosodic modifications (Cristia, 2013; Ma et al., 2011). Since men and women use such cues differently, understanding multimodal interaction can inform gender-tailored strategies for engaging with children. Third, speakers may distribute communicative load differently across modalities (e.g., Zhang et al., 2021; Zhou & Gu, 2024), with potential tradeoffs between gesture and speech (e.g., de Ruiter, 2006; Melinger & Levelt, 2004; Van der Sluis & Krahmer, 2007). For instance, when gestures are harder to produce, speakers may compensate by relying more on speech (Melinger & Kita, 2007). This trade-off may also apply to audience design in CDL, as certain signal channels may be less flexible, requiring adjustments in other channels to compensate. For example, Mandarin female speakers did not vary their speaking rate but lengthened their gestures when addressing children (Zhang & Gu, 2023). Thus, investigating one cue in isolation may not capture the full patterns of multimodal recipient design (Zhang et al., 2021). A multimodal approach can uncover these cross-modal interactions and compensatory mechanisms.

Given verbal and non-verbal differences between men and women mentioned above, it is possible that female and male speakers vary in their multimodal recipient design with different possible outcomes. For example, if women have greater adaptation in linguistic features to children, it may also exist in their non-verbal skills such as gestures. Alternatively, if men make lesser use of non-verbal features, they may compensate through other modalities like lexicon or prosody. The current study investigates these differences in the context of broadcasting language, where male and female broadcasters communicate child-directed and adult-directed content through speech, prosody, and gestures.

Broadcasting provides a unique domain for studying CDL, as children are exposed to information presented by both male and female broadcasters through television programmes. Prior research has found gender differences in news broadcasting, such as source selection (e.g., Craft & Wanta, 2004; Freedman & Fico, 2005), news topics (e.g., Ross & Carter, 2011), audience perception of credibility (Armstrong & McAdams, 2009), and competence (Brann & Himes, 2010). However, few studies have examined how broadcasters adjust their multimodal language for child versus adult audiences, let alone gender differences in these adjustments.

Zhang and Gu (2023) found that female broadcasters used a higher mean pitch, shorter utterances, more frequent words, and more pointing and representational gestures, but fewer pragmatic gestures when addressing children compared to adults. They also observed that female broadcasters did not slow their speech for children but extended the duration of their representational gestures. Whether male broadcasters exhibit similar or different adjustments and whether gender differences exist in the trade-off between modalities remain open questions.

Thus, this study addresses two main questions: (1) Do male broadcasters, like their female counterparts, adjust their lexical features (e.g., MLU, lexical diversity, word frequency, use of questions), prosody (e.g., pitch, intensity and speaking rate) and gestures (e.g., frequency, types and saliency) to children versus adult audiences? (2) Are there any gender differences in the degree of multimodal adjustments between hosting the child-directed and adult-directed TV programmes?

Method

Participants

Assuming a medium effect size (Cohen's d = 0.5), a significance level of $\alpha = 0.05$, and a desired power of 0.80, the analysis indicated that at least 32 participants per group would be necessary to reliably detect gender differences. In our study, both female (N=46, Mean age=19.6yrs, SD=0.91) and male (N=36, Mean age=20.6yrs, SD=0.55) Mandarin Chinese-speaking college students majoring in Chinese broadcasting from different grades (freshmen to senior years, Age range=18-23) took part in the study. Although the number of male and female participants was not perfectly balanced since there was an unequal number of females and males in the broadcasting major, the sample sizes within each gender group were sufficient to ensure adequate power for detecting medium-sized effects.

Procedures

Participants stood in front of a camera and were instructed to deliver news content as they would in an actual broadcast setting. Each participant delivered two types of programmes: adult-directed broadcasting (ADB) programmes and child-directed broadcasting (CDB) programmes. In both programmes, the same set of news topics was visually depicted with four corresponding pictures displayed on slides (see Figure 1 and pictures in Zhang & Gu, 2023). The sequence of programmes and pictures was counterbalanced across participants. Participants had two minutes to get familiar with the pictures and prepare their speech, and then they broadcasted for about one minute on each picture. To elicit natural speech and behaviours, participants were not specifically instructed to modify their speech or gestures; they were only asked to deliver the two types of programmes with an interval for a lunch break.

Figure 1: Setting of experiment.

Annotation and data processing

Speech was automatically transcribed via Xunfeitingjian (https://www.iflyrec.com/) and manually corrected. In total, we had 824.1 mins of recordings (female: 449.4 mins; male: 374.7 mins) and 19676 gestures (female: 10855, male: 8821).

Lexical/syntactic features We semi-automatically extracted lexical frequency (based on the SUBTLEX-CH corpus; Cai & Brysbaert, 2010), lexical diversity (based on moving type token ratio (MATTR), Covington & McFall, 2010), mean length of utterances (MLU, the average number of words per utterance; Dickinson & Porche, 2011) and the use of questions (proportion of questions in all sentences).

Prosody Utterance boundaries were annotated through the silence function in Praat (Boersma & Weenink, 2019) and then corrected manually. Mean pitch (semitone, ST), mean intensity (dB), speaking rate (the average number of words per second excluding pauses over 200ms; log-transformed, Han, 2019) and pauses (the number of pauses and pausing rate of broadcasting each picture) were coded and computed.

Gestures Using Elan (Wittenburg et al., 2006), the first author coded gestures in all clips and several second coders coded 16% of clips for verification, achieving a 92.1% agreement and a Kappa coefficient of 0.91, indicating substantial consistency. We categorized gestures into referential gestures and non-referential gestures (McNeill, 1992). Referential gestures can illustrate aspects of the topic being spoken of (such as representational gestures and pointing gestures), while beat gestures and pragmatic gestures (such as palm-up open gestures) are not referential but rather serve a discourse-pragmatic function (Kendon, 2017; Lopez-Ozieblo, 2020; McNeill, 1992). Additionally, we coded the saliency of each gesture based on the size of gestures in the manual articulators (part of the hands/arms that are being used: fingers, hands, arms), vertical and horizontal planes (see details in Chu et al., 2014; Zhang & Gu, 2023). We computed the frequency of referential and non-referential gestures per sec. and overall gesture saliency.

Analyses

All analyses were conducted in RStudio 9.0.351 (RStudio Team, 2020). To study the gender differences between ADB and CDB, we looked at lexical/syntactic (e.g., lexical diversity, word frequency, MLU, question frequency), prosodic (e.g., F0, intensity, speaking rate, pausing), and gestural (e.g., gesture saliency, frequency) features (DVs), and the effect of gender and its interaction with programme types (IVs). The four pictures and participants' studying grade (cohort) were control variables. Participants were included as a grouping variable with a random intercept. We constructed maximal mixed-effects models in the first instance. When the maximal models did not converge, we ran the models with different optimizers that allowed convergence or reduced the model complexity by removing the correlation between slope and intercept or random slope of interaction.

The codes can be viewed via the OSF link: https://osf.io/z8h37/?view_only=fc3cf43bbc8943fd86131575c88d9626.

Results

Table 1. presents the descriptive of multimodal features of female and male broadcasters in ADB and CDB.

Table 1: Means (SD) of multimodal features for programmes.

Dep variables	Programme	Female	Male
Lexical diversity	ADB	0.88 (0.03)	0.87 (0.03)
	CDB	0.89 (0.03)	0.88 (0.03)
MLU	ADB	26.2 (4.72)	29.4 (6.01)
	CDB	24.2 (4.33)	26.4 (5.68)
Word frequency	ADB	3.34 (0.09)	3.34 (0.08)
	CDB	3.37 (0.07)	3.39 (0.07)
Question	ADB	0.05 (0.07)	0.06 (0.09)
	CDB	0.14 (0.13)	0.12 (0.14)
Mean pitch	ADB	27.2 (1.26)	16.8 (1.71)
	CDB	28.0 (1.29)	17.7 (2.03)
Mean intensity	ADB	58.1 (3.08)	67.5 (2.66)
	CDB	58.3 (3.26)	68.2 (3.20)
Speaking rate	ADB	0.71 (0.03)	0.74 (0.04)
	CDB	0.72 (0.03)	0.75 (0.04)
Pause rate	ADB	0.33 (0.07)	0.44 (0.08)
	CDB	0.34 (0.07)	0.43 (0.08)
Gesture saliency	ADB	6.2 (1.31)	6.7 (0.63)
	CDB	6.5 (1.62)	6.8 (0.70)
Ref. gesture rate	ADB	0.13 (0.09)	0.12 (0.07)
	CDB	0.18 (0.10)	0.14 (0.08)
Non-ref. gesture rate	ADB	0.27 (0.13)	0.26 (0.15)
	CDB	0.23 (0.12)	0.27 (0.14)

Lexical/syntactic features

Lexical diversity Although programme types did not affect lexical diversity (β=0.004, p=0.16, 95%CI [-0.001, 0.009]), male broadcasters had less diverse vocabulary (β=-0.009, p=0.008, 95%CI [-0.02, -0.003], Figure 2A) than female ones.

MLU Both programme types and gender showed significant main effects on MLU. Male (β=-2.9, p<0.001, 95%CI [-4.03, -1.80]) and female broadcasters (β=-2.0, p<0.001, 95%CI [-2.99, -1.02]) all shortened their utterances in CDB compared to ADB. Additionally, males had longer utterances than their female counterparts (β=3.29, p=0.001, 95%CI [1.40, 5.17], Figure 2B).

Word frequency The results revealed a significant main effect on programme types, suggesting that both male broadcasters (β=0.05, p<0.001, 95%CI [0.04, 0.07]) and female broadcasters (β=0.03, p<0.001, 95%CI [0.01, 0.04]) used more frequent words to children than adults. However, no significant gender differences were observed, neither in ADB (β=0.02, p=0.10, 95%CI [-0.003, 0.04]) nor in CDB (β=-0.001, p=0.92, 95%CI [-0.02, 0.02]). Importantly, there was a significant interaction between gender and programme types (β=0.02, p=0.03, 95%CI [0.001, 0.04]), indicating that males ($Mean_{CDB\text{-}ADB}$=0.05) had more adjustments to children than females ($Mean_{CDB\text{-}ADB}$=0.03) did (see Figure 2C).

Proportion of questions The significant effect on programme types showed that both male broadcasters (β=0.06, p<0.001, 95%CI [0.04, 0.08]) and female broadcasters (β=0.09, p<0.001, 95%CI [0.07, 0.11]) used more questions in CDB than in ADB. Although gender did not affect the use of questions (β=0.008, p=0.63, 95%CI [-0.02, 0.04]), there was a significant interaction between gender and programmes (β=-0.03, p=0.03, 95%CI [-0.02, 0.04]), which means females had larger adjustments for children ($Mean_{CDB\text{-}ADB}$=0.09) than males ($Mean_{CDB\text{-}ADB}$=0.06) (see Figure 2D).

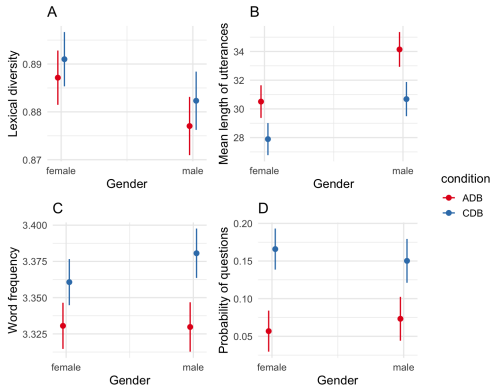

Figure 2: Predicted gender differences in linguistic/syntactic features (A: lexical diversity, B: MLU, C: word frequency, D: questions) between ADB and CDB.

Prosody

Mean pitch The significant positive effect of programme types revealed that male broadcasters (β=0.87, p<0.001, 95%CI [0.71, 1.04]) and female broadcasters (β=0.81, p<0.001, 95%CI [0.67, 0.96]) all raised their mean pitch for children than for adults. Additionally, gender also predicted mean pitch (β=-10.4, p<0.001, 95%CI [-11.1, -9.68]), that is, male broadcasters had lower mean pitch than females.

Mean intensity Both males (β=0.70, p<0.001, 95%CI [0.45, 0.95]) and females (β=0.28, p=0.01, 95%CI [0.06, 0.49]) had higher intensity when addressing children. Although male broadcasters had higher mean intensity than females (β=9.46, p<0.001, 95%CI [8.11, 10.80]), the further significant interaction between gender and programmes (β=0.42, p=0.012, 95%CI [0.09, 0.75]) suggested that male broadcasters had a larger adjustment to children than female broadcasters ($Male_{CDB\text{-}ADB}$=0.7dB vs. $Female_{CDB\text{-}ADB}$=0.2dB, see Figure3B).

Speaking rate and pause rate Programme types did not predict speaking rate (β=0.001, p=0.41, 95%CI [-0.002, 0.005]) or pause rate (β=0.0004, p=0.95, 95%CI [-0.01, 0.01]), but male broadcasters spoke faster (β=0.03, p<0.001, 95%CI [0.02, 0.05]) and paused more frequently (β=0.10, p<0.001, 95%CI [0.07, 0.13]) than females. Additionally, there was no interaction between gender and programme for speaking rate (β=0.0003, p=0.89, 95%CI [0.07, 0.13]) or pause rate (β=-0.005, p=0.65, 95%CI [-0.03, 0.02]).

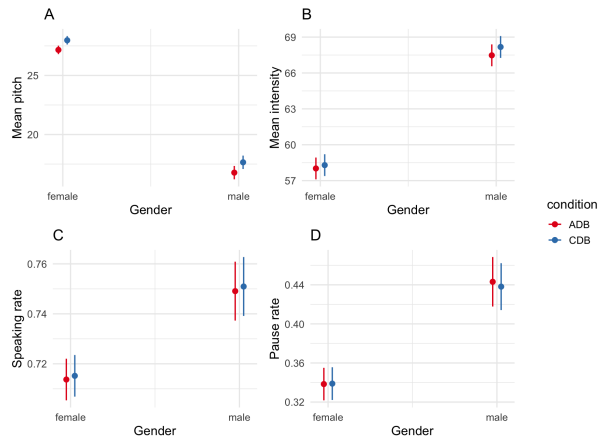

Figure 3: Predicted gender differences in prosodic features (A: mean pitch, B: mean intensity, C: speaking rate, D: pause rate) between ADB and CDB.

Gestures

Gesture saliency Only female broadcasters (β=0.25, p<0.001, 95%CI [0.14, 0.36]) but not male broadcasters (β=0.06, p=0.37, 95%CI [-0.07, 0.18]) had more salient gestures for children than for adults. However, male broadcasters generally had more salient gestures than their female counterparts (β=0.50, p=0.01, 95%CI [0.13, 0.88], see Figure 4 for the comparison of gesture saliency). Moreover, there was a significant interaction between gender and condition (β=-0.19, p=0.03, 95%CI [-0.36, -0.02]), indicating that females ($Mean_{CDB\text{-}ADB}$=0.3) had more adjustments to children than males ($Mean_{CDB\text{-}ADB}$=0.1) did (see Figure 5A).

Figure 4: One female broadcaster (above) and one male broadcaster (below) indicating lifting a dog in ADB (left) and CDB (right).

Referential gesture rate Both male (β=0.02, p=0.03, 95%CI [0.03, 0.06]) and female broadcasters (β=0.04, p<0.001, 95%CI [0.008, 0.04]) produced more frequent referential gestures for children than for adults. Additionally, male broadcasters produced fewer referential gestures than female broadcasters did in general (β=-0.01, p=0.01, 95%CI [-0.03, -0.004]). Moreover, as shown in Figure 5B, there was a significant interaction between gender and programmes (β=-0.01, p=0.06, 95%CI [-0.04, 0.001]), indicating that females ($Mean_{CDB\text{-}ADB}$=0.05) had larger adjustments in referential gestures for children than males ($Mean_{CDB\text{-}ADB}$=0.02).

Non-referential gesture rate Only females made fewer non-referential gestures in CDB than in ADB (β=-0.04, p<0.001, 95%CI [-0.06, -0.02]) but male broadcasters did not make significant differences (β=0.006, p=0.62, 95%CI [-0.06, -0.02]). Such gender differences in adjustments were also confirmed by a significant interaction between gender and programmes (β=0.05, p=0.005, 95%CI [0.01, 0.08]), where females had larger adjustments ($Mean_{CDB\text{-}ADB}$=-0.04) between the two programmes than males ($Mean_{CDB\text{-}ADB}$=0.01) did (Figure 5C).

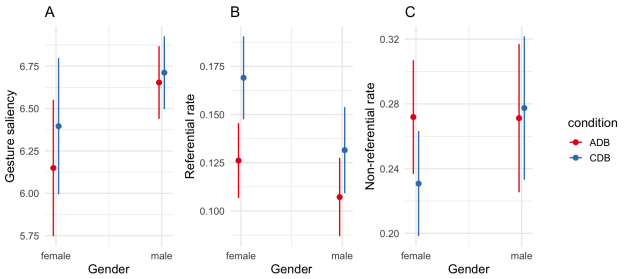

Figure 5: Predicted gender differences in gestures (A: gesture saliency; B: referential gesture rate; C: non-referential gesture rate) between ADB and CDB.

Discussion

This study investigated gender differences in child-directed and adult-directed multimodal communication, focusing on lexical/syntactic, prosodic, and gestural aspects of multimodal language production. The results indicate that while both male and female broadcasters adjust their communication strategies for child audiences, they exhibit distinct patterns across different modalities and conditions (see Table 2 for a summary of the main gender differences).

Table 2: Gender differences in multimodal features.

Dep variables	M vs. F	Dep variables	M vs. F
Lexical diversity	<	Speaking rate	>
MLU	>	Pause rate	>
Word frequency	NS.	Saliency	>
Question	>	Ref. gesture rate	<
Mean pitch	<	Non-re. gesture rate	NS.
Mean intensity	>		

Note: M vs. F means male vs. female, ＞means males adjust more than females, ＜ means males adjust less than females, NS means no significant difference.

Multimodal audience design in specific context

Both male and female broadcasters adjusted most multimodal cues for children, except for speaking rate, pause rate and lexical diversity. Interestingly, neither gender slowed their speech for child audiences, contrary to previous studies that CDL is typically slower than ADL (Golinkoff et al., 2015; Soderstrom, 2007). However, this pattern may vary across languages and contexts. For example, Mandarin CDL does not consistently exhibit a slower speaking rate and in some cases, can even be faster than that of ADL for long utterances (Han, de Jong, & Kager, 2021; Han, Yang, & Gu, 2024). Additionally, broadcasters did not pause more frequently in CDB, which may be attributed to the one-minute time constraint in our experiment. To complete their speech within the time limit, broadcasters may have sacrificed frequent pausing and slower speaking rates in both programmes. Thus, whether Mandarin CDL slows down may depend on linguistic context and child age.

Furthermore, while CDL is usually simpler than ADL, broadcasters did not significantly reduce their lexical diversity for children. For TV programmes, the context requires them to be more formal, and broadcasters have time to prepare in advance to maintain coherence and structure, in contrast to spontaneous, daily conversation.

Gender differences in adaptation

Despite general adjustments in child-directed programmes, notable gender differences remained (Table 2).

Females had larger adjustments in non-verbal expressions compared to males. Although males exhibited more salient gestures than females, yet only females adjusted their gesture saliency to children compared to adults. It had been found that men use more expansive gestures than women (Hall, 1984; Wolfe, 2005). Additionally, men's larger physical size may also contribute to the perceived expansiveness of their gestures. According to the criteria we classify the size of a gesture, generally, a man will extend his gestures over a larger area than a woman because of their longer arms. However, while males tended to use larger gestures regardless of audiences, females extended their gestures especially for children, making their non-verbal expression

more exaggerated and easily noticeable, highlighting their greater adaptation for children not only in more salient prosody (Han et al., 2022) but also in more salient non-verbal communication.

In addition, our study further revealed that females adjusted more in the use of meaningful gestures. Although both genders increased their use of referential gestures in CDB, females adjusted the frequency of referential gestures more than males. Referential gestures, which directly link speech to visual cues, facilitate children's comprehension, whereas non-referential gestures are more abstract and may be harder for children to interpret. However, male broadcasters did not significantly modify their use of non-referential gestures when addressing children. In contrast, females used significantly fewer non-referential gestures than males in CDB, suggesting that women may be more effective in optimizing gestures for child audiences.

Females also demonstrated larger adaptations than males in terms of lexical/syntactic features, such as using more diverse words and shorter utterances. Despite the general tendency for CDL to feature less diverse words (e.g., Tal & Arnon, 2018), female broadcasters used more diverse words than males, likely to maximize informational content within the time constraints. Moreover, female broadcasters used shorter utterances to both audiences, likely to enhance clarity and comprehension. Under the same one-minute broadcast constraints, males, on the other hand, had longer utterances and lower lexical diversity, indicating the differences in language strategies between males and females. Additionally, although both broadcasters used more questions to children, females utilised more questions to children than to adults comparing to males. Admittedly, during adult-directed programmes, broadcasters may focus more on content delivery, yet in CDB, questions serve to engage young audiences and create a communicative, semi-interactive atmosphere. Female broadcasters used more questions to encourage active thinking, which may indicate stronger communicative skills with children.

There are also gender differences in prosodic features, with males having higher mean intensity and faster speaking rate with more frequent pausing than females. Some prior research has suggested that males generally speak faster than females in Chinese, English, and Irish English (Fitzsimmons et al., 2001; Lee & Doherty, 2017; Yuan, Liberman, & Cieri, 2006), though some studies reported no significant differences in American and Australian English (Block & Killen, 1996; Robb et al., 2004). In this study, males indeed had a faster speaking rate and more frequent pauses in Mandarin than females. Given that nervousness can increase the speed of an individual's speech (Hughes, Harrison & M. de Haan, 2020; Kasl & Mahl, 1965), and fluency is associated with fewer and shorter pauses (de Jong, 2018), the slower speaking rate observed in males may suggest that they have greater difficulty in conceptualizing or processing their speech in this task, though further research is needed to determine whether this reflects a general cognitive difference or a task-specific effect.

It is not surprising that both males and females raise their mean pitch and mean intensity for children, but females had relatively higher mean pitch than males, while males had higher mean intensity than females in both programmes. This aligns with physiological differences: males have a greater physiological intensity range (e.g., Coleman et al., 1977), resulting in the higher vocal intensity, while females have greater vocal fold stiffness, contributing to higher pitch (e.g., Kent, 1976; Titze, 1988, 1989).

Compensation through modalities

The observed gender differences may emerge due to evolutionary and neurological factors that enhance women's sensitivity in communication, particularly in child-related contexts (de Pisapia et al., 2013; Geary, 2000). These findings extend our understanding of multimodal recipient design across different genders. While broadcasters adjust their multimodal language based on recipients, the direction or the degree of adjustments differs between females and males. Females excel in certain multimodal aspects, such as using more meaningful gestures and producing shorter utterances and more diverse words when talking to children. Meanwhile, males have longer utterances, more pausing that they may compensate for through other modalities: Compared with females, they used more frequent words (which are easier to understand), higher vocal intensity (which helps capture attention), and faster speaking rates (which create a more energetic tone) for children, ensuring their communication remains effective for child audiences. These compensatory strategies suggest that gendered variations in communication extend beyond physical and psychological factors, reflecting adaptive mechanisms across different modalities. However, it remains to be seen how these gendered communication strategies differ across cultures and affect audience perception.

Conclusion

This study provides new insights into gender differences in multimodal communication and audience design, specifically in the context of child-directed and adult-directed broadcasting. These findings enhance existing research on CDL by offering a comprehensive perspective on lexical, syntactic, prosodic, and gestural dimensions of multimodal communication, thereby contributing valuable insights to both child developmental and gender studies. Moreover, understanding these gender differences not only enriches our knowledge of communication dynamics but also informs practical applications. By identifying the types of programmes in which men and women excel or appear to be deficient, we can develop more tailored broadcasting training that capitalizes on their unique strengths. Furthermore, for audiences, targeted language production can lead to improved comprehension of the content presented in these programmes. It remains to be seen which kinds of adjustments are more favoured by children (e.g., more referential gestures, more frequent words, a faster speaking rate, etc.) and whether gender differences would affect children's likability and remembering of contents.

References

Armstrong, C., & McAdams, M. (2009). Blogs of information: How gender cues and individual motivations influence perceptions of credibility. *Journal of Computer-Mediated Communication*, *14*(3), 435-456.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, *15*, 415-419.

Benders, T., StGeorge, J., & Fletcher, R. (2021). Infant-directed speech by Dutch fathers: Increased pitch variability within and across utterances. *Language Learning and Development*, *17*(3), 292-325.

Block, S., & Killen, D. (1996). Speech rates of Australian English-speaking children and adults. *Australian Journal of Human Communication Disorders*, *24*, 39-44.

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* (6.1.08) [Computer software].

Brann, M., & Himes, K. (2010). Perceived credibility of male versus female television newscasters. *Communication Research Reports*, *27*(3), 243-252.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS ONE*, *5*(6), e10729.

Cameron, H., & Xu. X. (2011). Representational gesture, pointing gesture, and memory recall of preschool children. *Journal of Nonverbal Behavior*, *35*, 155-171.

Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, *47*(1), 14-27.

Cebioglu, S, Marin, K. A., & Broesch, T. (2022). Variation in caregivers' references to their toddlers: child-directed speech in Vanuatu and Canada. *Child Development*, *93*(6), 622-638.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694-709.

Coleman, R. F., Mabis, J. H., & Hinson, J. K. (1977). Fundamental frequency–sound pressure level profiles of adult male and female voices. *Journal of Speech and Hearing Research*, *20*, 197-204.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics*, *17*(2), 94-100.

Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2022). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 1-20.

Craft, S., & Wanta, W. (2004). Women in the newsroom: Influences of female editors and reporters on the news agenda. *Journalism & Mass Communication Quarterly*, *81*(1), 124-38.

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass*, *7*(3), 157-170.

Davidson, R. G., & Snow, C. E. (1996). Five-year-olds' interactions with fathers versus mothers. *First Language*, *16*, 223-242.

de Pisapia, N., Bornstein, M. H., Rigo, P., Esposito, G., de Falco, S., & Venuti, P. (2013). Sex differences in directional brain responses to infant hunger cries. *NeuroReport*, *24*(3), 142-146.

de Ruiter, J. P. (2006). Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech Language Pathology*, *8*, 124-127.

Dickinson, D. K., & Porche, M. V. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development*, *82*(3), 870-886.

Dong, S., Gu, Y., & Vigliocco, G. (2021). The impact of child-directed language on children's lexical development. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43. Retrieved from https://escholarship.org/uc/item/38x9h9h4

Donnellan, E., Jordan-Barros, A., Theofilogiannakou, N., Brekelmans, G., Murgiano, M., Motamedi, Y., Grzyb, B., Gu, Y., & Vigliocco, G. (2023). The impact of caregivers' multimodal behaviours on children's word learning: A corpus-based investigation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45. Retrieved from https://escholarship.org/uc/item/6km748xv

Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 345-382). Oxford University Press.

Fitzsimmons, M., Sheahan, N., & Staunton, H. (2001). Gender and the integration of acoustic dimensions of prosody: Implications for clinical studies. *Brain & Language*, *78*(1), 84-108.

Freedman, E., & Fico, F. (2005). Male and female sources in newspaper coverage of male and female candidates in open races for governor in 2002. *Mass Communication and Society*, *8*(3), 257-272.

Gallaway, C., & Richards, B. J. (1994). *Input and Interaction in Language Acquisition*. Melbourne: Cambridge Univ. Press.

Gangstead, S. W., & Simpson, J. A. (2000). The evolution of human mating: Tradeoffs and strategic pluralism. *Behavioral and Brain Sciences*, *23*, 573–644.

Geary, D. C. (2000). Evolution and proximate expression of human parental investment. *Psychological Bulletin*, *126*, 55-77.

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me the social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344.

Grinberg, D., Levin-Asher, B., & Segal, O. (2022). The myth of women's advantage in using child-directed speech: Evidence of women versus men in single-sex–parent families. *Journal of Speech, Language, and Hearing Research*, *65*(11), 4205-4227.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore: Johns Hopkins University Press.

Han, M. (2019). *The role of prosodic input in word learning: A cross-linguistic investigation of Dutch and mandarin Chinese infant-directed speech*. Amsterdam: LOT Publisher.

Han, M., de Jong, N. H., & Kager, R. (2021). Language specificity of infant-directed speech: Speaking rate and word position in word-learning contexts. *Language Learning and Development*, *17*(3), 221-240.

Han, M., de Jong, N. H., & Kager, R. (2022). Prosodic input and children's word learning in infant- and adult-directed speech. *Infant Behavior and Development*, *68*, 101728.

Han, M., de Jong, N. H., & Kager, R. (2023). Relating the prosody of infant-directed speech to children's vocabulary size. *Journal of Child Language*. *2023*, 1-17.

Han, M., Yang, L., & Gu, Y. (2024). Faster and smoother: Fluency in Chinese child-directed speech. *Speech Prosody 2024*.

Hagoort, P., & Özyürek, A. (2024). Extending the architecture of language from a multimodal perspective. *Topics in Cognitive Science*. Advance online publication.

Jacobson, J. L., Boersma, D. C., Fields, R. B., Olson, K. L., & David, C. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, *54*(2), 436-442.

Hughes S. M., Harrison M. A., & de Haan K. M. (2020). Perceived nervous reactions during initial attraction and their potential adaptive value. *Adaptive Human Behavior and Physiology*, *6*(1), 30-56.

Kasl, S. V., & Mahl, G. F. (1965). The relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology*, *1*(5), 425-433.

Kavakli, M., & Chen, Y.L. (2014). Gender Differences in Gesture based Interaction. *Proceedings of the Seventh International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services*, 15-21.

Kempe, V. (2009). Child-directed speech prosody in adolescents: Relationship to 2D:4D, empathy, and attitudes towards children. *Personality and Individual Differences*, *47*(6), 610-615.

Kendon, A. (2017). Pragmatic functions of gestures. *Gesture*, *16*(2), 157-175.

Kent, R. (1976). Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. *Journal of Speech and Hearing Research*, *19*, 421-447.

Kondaurova, M. V., VanDam, M., Zheng, Q., & Welikson, B. (2023). Fathers' unmodulated prosody in child-directed speech. *Journal of the Acoustical Society of America*, *154*(6), 3556-3567.

Kristensen, N. M., Sundby, C. F., Hauge, M. N., & Lofkvist, U. (2020). Female caregivers talk more to 18–56-months-old children with and without hearing impairment than male caregivers measured with LENATM – A cross-sectional pilot study. I*nternational Journal of Pediatric Otorhinolaryngology*, *130*, 109809.

Lee, A., & Doherty, R. (2017). Speaking rate and articulation rate of native speakers of Irish English. *Speech, Language and Hearing*, *20*(4), 206-211.

Leerkes, E. M., Blankson, A. N., & O'Brien, M. (2009). Differential effects of maternal sensitivity to infant distress and nondistress on social-emotional functioning. *Child Development*, *80*, 762-775.

Lopez-Ozieblo, R. (2020). Proposing a revised functional classification of pragmatic gestures. *Lingua*, 102870.

Majorano, M., Rainieri, C., & Corsano, P. (2013). Parents' childdirected communication and child language development: A longitudinal study with Italian toddlers. *Journal of Child Language*, *40*(4), 836-859.

Malone, M. J., & Guy, R. F. (1982). A comparison of mothers' and fathers' speech to their 3-year-old sons. *Journal of Psycholinguistic Research*, *11*(6), 599-608.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.

McGregor, K. K. (2008). Gesture supports children's word learning. *International Journal of Speech-Language Pathology*, *10*(3), 112-117.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, *22*, 473–500.

Melinger, A., & Levelt, W. J. M. (2004). Gesture and the communicative intention of the speaker. *Gesture*, *4*, 119–141.

Nicoladis, E., Pika, S., Yin, H., & Marentette, P. (2007). Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, *28*, 721-735.

Pancsofar, N., & Vernon-Feagans, L. (2006). Mother and father language input to young children: Contributions to later language development. *Journal of Applied Developmental Psychology*, *27*(6), 571-587.

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Robb, M. P., Maclagan, M. A., & Chen, Y. (2004). Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics & Phonetics*, *18*(1), 1-15.

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversation exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700-711.

Rondal, J. A. (1980). Fathers' and mothers' speech in early language development. *Journal of Child Language*, *7*(2), 353-369.

Ross, K., & Carter, C. (2011). Women and news: A long and winding road. *Media, Culture and Society*, *33*(8), 1148-1165.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1), 185-205.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*, 1762-1774.

Rowe, M. L., Coker, D., & Pan, B. A. (2004). A comparison of fathers' and mothers' talk to toddlers in low-income families. *Social Development*, *13*(2), 278-291.

Shi, J., Gu, Y., & Vigliocco, G. (2022). Prosodic modulations in child-directed language and their impact on word learning. *Developmental Science*, e13357.

Shneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659-673.

Snow, C. (1995). Issues in the study of input: Finetuning, universality, individual and developmental differences, and necessary causes. *In P. Fletcher & B. MacWhinney (Eds.), The handbook of child language* (pp. 180-193). Oxford: Blackwell Publishers.

So, W. C., Sim, C., & Low W.S. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665-681.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532.

Spinelli, M., & Mesman, J. (2018). The regulation of infant negative emotions: The role of maternal sensitivity and infant-directed speech prosody. *Infancy*, *23*(4), 502-518.

Tal, S., & Arnon, I. (2018). SES effects on the use of variation sets in child-directed speech. J*ournal of Child Language*, *45*(6), 1423-1438.

Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *The Journal of Acoustic Society of America*, *83*, 1536-1552.

Titze, I. R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of Acoustic Society of America*, *85*, 1699-1707.

VanDam, M., Thompson, L., Wilson-Fowler, E., Campanella, S., Wolfenstein, K., & De Palma, P. (2022). Conversation initiation of mothers, fathers, and toddlers in their natural home environment. *Computer Speech & Language*, *73*, 101338.

Van der Sluis, I., & Krahmer, E. (2007). Generating multimodal references. *Discourse Processes*, *44*, 145-174.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. *Proceedings of the 5th International Conference on Language Resources and Evaluation*, 1556-1559.

Wolfe, J. (2005). Gesture and collaborative planning: A case study of a student writing group. *Written Communication*, *22*(3), 298-332.

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. *Ninth international conference on spoken language processing*.

Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J. I., & Vigliocco, G. (2021). More than words: Word predictability, prosody, gesture and mouth movements in natural language comprehension. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1955), 20210500.

Zhang, Y., & Gu, Y. (2023). A recipient design in multimodal language on TV: A comparison of child-directed and adult-directed broadcasting. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 45. Retrieved from https://escholarship.org/uc/item/17k7h7m6

Zhang, Y., & Gu, Y. (2024). Individual differences in multimodal child-directed language: Unraveling individual styles, empathy and the Big Five personality traits. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46. Retrieved from https://escholarship.org/uc/item/922774vn

Zhou, J., & Gu, Y. (2024). The Impact of Teachers' Multimodal Cues on Students' L2 Vocabulary Learning in Naturalistic Classroom Teaching. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 46. Retrieved from https://escholarship.org/uc/item/8jg192cm