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

Title
A recipient design in multimodal language on TV: A comparison of child-directed and adult-directed broadcasting

Permalink
https://escholarship.org/uc/item/17k7h7m6

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
zhang, yanran
Gu, Yan

Publication Date
2023

Peer reviewed

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

A recipient design in multimodal language on TV: A comparison of child-directed and adult-directed broadcasting

Yanran Zhang (yanran.zhang@cuz.edu.cn)
Communication University of Zhejiang

Yan Gu (yan.gu@essex.ac.uk)
Department of Psychology, University of Essex & Department of Experimental Psychology, UCL

Abstract

Child-directed language is a unique multimodal communication behaviour that differs from adult-directed language. We investigated how broadcasters organize their multimodal language production on an adult and child-directed programme to better understand the recipient design in the broadcasting context. Thirty-six future broadcasters produced live programmes for children and adults, respectively, whose linguistic features (utterance=3888), speech prosody, and gestures (N=8486) were analysed as a function of programme. We found that broadcasters used a higher mean pitch but a smaller pitch range, shorter utterances, high(er) frequency words, more questions, pointing and representational gestures but fewer pragmatic gestures in child-directed broadcasting. Gestures were also more salient and slower when addressing children audiences. However, there were no differences in lexical diversity, speaking rate, pausing, or beat gestures between programmes. In conclusion, broadcasters did engage in recipient design multimodally, but the distinction between the speaker and audience orientation is not binary but should be understood across signal channels according to contexts.

Keywords: recipient design; multimodal communication; child-directed language; broadcasting; gesture

Introduction

People are capable of communicating according to various situations. For example, they may adjust their speech, intonation, and gestures according to different addressees, such as when talking to children versus adults. The process of tailoring language for an addressee is called recipient design (Sacks et al., 1974) or audience design (Clark & Murphy, 1982). Child-directed language (CDL) is a typical recipient design where caregivers talk to their children at a higher pitch level and a wider pitch range with shorter and simpler utterances. CDL is found across languages (Cox et al., 2022; Fernald et al., 1989). Even in sign languages, caregivers sign slowly with exaggerated movements or frequently repeat the same sign (Masataka, 1992; Perniss et al., 2017). So far, studies investigating CDL have mainly focused on family dyads, but few have investigated the recipient design in vastly different situations, such as the broadcasting context. Particularly, broadcasters' speech is often considered more formal, faster, and higher in pitch than ordinary conversations (e.g., Cotter, 1993; Medrado et al., 2005), and there are usually only imagined addressees. It is entirely unknown how broadcasters use different signal channels (speech, intonation, gestures, etc.) as a communicative strategy in child-directed and adult-directed programmes, and whether they adopt recipient design in all information channels. This study compares future broadcasters' multimodal language production between their child-directed broadcasting (CDB) and adult-directed broadcasting (ADB).

Broadcasting language

Newsreaders have often served as role models within a specific language community (Swerts & Krahmer, 2010). Various aspects of broadcasting languages have already been examined (e.g., Ambrazaitis & House, 2017; Rodero & Cores-Sarría, 2021; Swerts & Marsi, 2012). For example, Rodero and Cores-Sarría (2021) found that the standard speaking style of newsreaders is characterized by specific intonation and rhythm (e.g., a higher pitch level and an upward and downward intonation). Ambrazaitis and House (2017) found that head beats and eyebrow movements were associated with the phonological prosodic structure in Swedish newsreaders. In addition, by analysing the facial expressions and speech of two Dutch newsreaders (children's versus adults' programme), Swerts and Krahmer (2010) found that newsreaders strive to make their facial expressions congruent with verbal content that are more expressive when addressing children than adults in general.

Although a handful of studies have attempted to clarify characteristics in the broadcasting context, they used a between-subject comparison and had a small sample size (e.g., Mok, Fung & Li, 2014; Swerts & Krahmer, 2010). For example, Mok et al. (2014) studied the prosodic features of Cantonese newsreaders and non-newsreader speakers, but the contents of the news announced by each anchor were different, and the control group had one extra day to prepare. Swerts and Krahmer (2010) first attempted to investigate multiple cues in broadcasting programmes, but there were only two newsreaders. There is a paucity of research on social media, such as YouTube and its possible effects on young children's learning and development (Izci et al., 2019; Neumann & Herodotou, 2020). Also, researchers encouraged broadcasters to consider children audiences' characteristics when organizing children's programmes (Abelman & Atkin, 2000), as television programmes are related to children's physical health, physical activities, interest in the study, etc. (Jenvey, 2007; Gupta et al., 1994). However, it is still unclear how the same broadcasters adjust their multimodal

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.* ©2023 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

communication when producing the same content for adults and children in a naturalistic setting.

Child-directed and Adult-directed language

CDL differs from ADL in many aspects, and it facilitates children's language learning and acquisition (e.g., Anderson et al., 2021; Golinkoff et al., 2015; Shi, Gu, & Vigliocco, 2022; Spinelli & Mesman, 2018). For example, CDL usually has exaggerated intonation and more varying pitch, a slower speaking rate, a shorter utterance length, and lower lexicon diversity with more iconic gestures (e.g., Campisi & Özyürek, 2013; Cristia, 2013; Han, De Jong & Kager, 2022), and English caregivers' use of a higher mean pitch correlates with their children's immediate unknown word learning outcome (Shi et al., 2022). However, most CDL studies have investigated communication between familiar acquaintances, such as mother/father/grandmother/sibling-child interaction (e.g., Cristia, 2013; Hoff & Krueger, 1991; Lamb & Lewis, 2010; Shute & Wheldall, 1999, 2001), or caregivers in schools (e.g., Kempe, 2009). Furthermore, CDL research typically focuses on characterising a few specific features (e.g., speaking rate, pitch, syntax, etc.) (Raneri et al., 2020; Song, Demuth, & Morgan, 2010; Zangl et al., 2005), and the age of children in such studies is usually below five years old.

So far, rare research has investigated CDL and ADL in the broadcasting context, which is different from typical caregiver-child dyads in terms of addressees and children's ages. Given the global picture of children being increasingly exposed to social media, which is now an essential part of everyone's life, broadcasting language is an important area for research into child-directed communication. However, we know little about how TV broadcasters use multimodal language to communicate with children and adults.

Multimodality in recipient design

Engaging in recipient design affects speech production (Snow & Ferguson, 1977; Isaacs & Clark, 1987; Galati & Brennan, 2010). For example, Tippenhauer et al. (2020) examined audience design in CDL and ADL, showing that speakers reduced word duration less in CDL than in ADL. Recipient design also shapes other communication means, such as co-speech gestures. Tippenhauer et al. (2020) only looked at recipient design in word duration, but the differences in gesture duration (the temporal length of a gesture) are unknown. Equivalent to the reasoning for word duration in linguistic aspects, the average duration of a gesture may also be shaped by recipient design when talking to different addressees.

Some aspects of co-speech gestures relate to mutually shared knowledge (common ground, Clark, 1996; Holler & Stevens, 2007; Holler & Wilkin, 2009). For example, speakers used more words and gestures when narrating novel contents and produced representational gestures at a higher rate, but significantly reduced the number of words and shortened repeated gestures in duration when common ground existed (e.g., Campisi & Ozyürek, 2013; Galati & Brennan, 2010; Holler & Stevens, 2007; Schubotz, Ozyürek, & Holler, 2018, 2022).

Also, researchers who study language production debate whether audience design is driven by speakers' needs or by their audiences' needs (e.g., Arnold et al., 2012; Aylett & Turk, 2004; Bard et al., 2000), but they always looked at specific features that cannot be generalized to different situations. For example, broadcasting is a unique situation where speakers do not have an addressee but seriously care about their imagined audiences. Hence, to contribute to this debate, we ask two detailed questions: First, how do TV broadcasters use multimodal language (e.g., speech, prosody, and gestures) when addressing different audiences (adults versus children)? Second, is multimodal broadcasting speaker-centred or listener-centred?

Method

Participant

Thirty-six participants majoring in Chinese broadcasting and hosting participated in the study (Mean age=19.7yrs, SD=0.93, all native Mandarin speakers). Several studies have found that mothers and fathers differentiated their language use when talking to their children (Van de Mieroop, Zenner, & Marzo, 2016). Also, mothers vocalized more frequently and for a longer time than fathers (Kiepura, Niedźwiecka, & Kmita, 2021). Moreover, in formal contexts that involved teaching and learning, the coda laterals of mothers were significantly darker than fathers (Sim, 2021). To avoid gender differences, in our study the participants were all female. Participants gave written informed consent to use the data before the recording started.

Stimuli

Participants were required to do a live broadcast explaining pictures in front of a camera both on a regular TV (adult-directed) and a children's (child-directed, to students at primary schools) programme (see Figure 1 for experiment setting). In total, there were four pictures chosen from life-related aspects that adults and children easily understood.

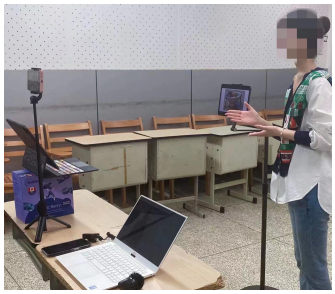

Figure 1: setting of the experiment.

Procedure

Both programmes were recorded on the same day with a lunch break in between (sequences counterbalanced). They

had two minutes to prepare and talked for around one minute for each picture. The sequence of pictures was also counterbalanced. The mean talking duration of each picture in the ADB (M=74.63 sec, SD=21.08) and CDB (M =73.68 sec, SD =26.53) did not differ significantly (β=-1.11, p=.59). The speech was recorded through Audacity (44.1 kHz, 16-bit rate) with a wireless clip-on. The video was recorded by a high-quality mobile phone with a resolution of 2772×1344. Gestures were not mentioned at any point in the experiment.

Coding

Linguistic features of speech

Speech was transcribed via a transcription website (https://www.iflyrec.com/zhuanwenzi.html), and errors were corrected manually. We measured:

(1) *Mean length of utterance* (MLU), the average number of words per utterance (Dickinson & Porche, 2011);

(2) *Lexical diversity*, measured by MATTR (moving average type-token ratio), which is less biased than TTR (type-token ratio) as it calculates average TTRs for successive sample pieces of a specified token size (Covington & McFall, 2010);

(3) *Sentence types* (statements and questions);

(4) *Word frequency*, calculated through the SUBTLEX-CH corpus (Cai & Brysbaert, 2010), which comprises Chinese word frequency based on Chinese subtitles.

Prosody

Boundaries of sentences were annotated in Praat (Boersma & Weenink, 2019). We measured:

(1) *Speaking rate* (the average number of words per second excluding pauses over 200ms (Han, 2019));

(2) *Pitch* (semitone, ST): mean F0, F0 SD and F0 range;

(3) *Intensity* (dB): mean intensity and intensity range;

(4) *Pauses* (the number of pauses, mean pausing duration and pausing rate of broadcasting each picture).

Gesture

Gestures were coded in ELAN (Wittenburg et al., 2006). We measured:

(1) *Gesture types*:

Representational gestures: metaphorical or iconic gestures to illustrate the concept by drawing its outline, indicating its shape or representing it (McNeill, 1998; Müller, 1998).

Emblem gestures: conventionalised gestures in certain areas (McNeill, 1992; Kendon, 2004). For example, a ring formed by the thumb and the index finger means OK in most countries, but it means "zero" in France (Morris et al., 1979). In actual coding, we merged emblems and representational gestures together.

Pointing gestures: a finger extends in the direction of something or points without any visible target, such as time or location (McNeill, 1992, 1998).

Beat gestures: simple motoric movements produced along with the rhythm of the speech (e.g., a hand with an open palm flipping outwards) (So, Sim, & Low, 2012). In actual coding, we coded repetitive beating movements as one beat gesture, but also coded the repeated times of every beat.

Pragmatic gestures: such as palm-up open gestures, serve a primarily or exclusively discourse-pragmatic function and lack an identifiable referent in the topic of talk (e.g., Kendon, 2017; Lopez-Ozieblo, 2020; McNeill, 1992).

(2) *Gesture rate* (the number of gestures per second);

(3) *Gesture duration* (the temporal length of each gesture);

(4) *Gesture saliency*. We coded saliency in three dimensions, one is the manual articulators (fingers, hand, forearm, or whole arm, depending on which part of the hand and/or arm was moving), one is the vertical zone (below the waist, between waist and chest, between chest and chin, and over the chin) (Chu et al., 2014), and thirdly the horizontal zone (narrower than the head, between head and shoulder, and broader than the shoulder). Each gesture was given 1 point if it was a finger movement/below the waist/narrower than the head and 1 additional point for each increment in each dimension.

Data analysis

In total, we had 351.61 minutes of audiovisual recordings. There were 3888 utterances (1891 in ADB and 1997 in CDB), and 8486 gestures (4339 ADB and 4147 CDB).

Linear mixed-effects models in the R environment (R Core Team, 2022) were used to assess differences in linguistic features, prosody, and gestures (dependent variables) as a function of broadcasting programmes (adult or child). Logistic regressions were used when dependent variables were binary (e.g., sentence types of question vs. statement). Participants were included as a grouping variable with a random intercept. The four pictures were control variables.

Result

Table 1 presents the results of linguistic cues, where broadcasters produced shorter utterances, words of a higher frequency, and a larger proportion of questions in CDB than in ADB (all *p*'s<.001). The results suggest that broadcasters structured their speech to be more easily understood. However, in terms of lexical diversity, there were no significant differences (*p*=.19) between the two programmes.

Table 1: Means (SD) and results (beta and *p*) of linguistic features for each condition.

Dep variables	ADB	CDB	*β and p*
MLU	26.05(4.97)	23.81(4.49)	-2.24***
Lexical diversity	0.88(0.03)	0.89(0.03)	4.03
Word frequency	3.34(0.08)	3.37(0.08)	3.37***
Question	0.69(1.04)	2.02(1.80)	1.13***

Note: ***$p < .001$, **$p < .01$, *$p < .05$.

For prosody (Table 2), broadcasters produced a higher mean pitch, a greater pitch variation, a larger pitch maximum and pitch minimum, and a smaller pitch range in CDB than ADB (all *p*'s<.001), indicating that broadcasters used a more exaggerated vowel to attract children's attention. However, there were no significant differences in the speaking rate

(p=.65), pausing rate (p=.36), mean pausing duration (p=.54), mean intensity (p=.12) and intensity range (p=.82).

Table 2: Means (SD) and results (beta and p) of prosodic cues for each condition.

Dep variables	ADB	CDB	β and p
Mean_F0(ST)	27.15(1.22)	27.99(1.29)	0.83***
SDF0 (ST)	58.47(6.39)	60.73(6.98)	2.31***
Max_F0 (ST)	36.07(1.52)	36.53(1.37)	0.47***
Min_F0 (ST)	16.14(2.07)	17.04(1.91)	0.89***
F0 range (ST)	15.08(2.59)	14.28(2.30)	-0.78***
Speaking rate	0.71(0.03)	0.71(0.04)	-0.002
Avg_Intensity (dB)	58.13(3.33)	58.30(3.57)	0.18
Intensity range (dB)	34.91(2.47)	34.88(2.55)	-0.034
Pausing rate	0.33(0.07)	0.33(0.07)	0.36
Avg. pause dur (sec)	0.40(0.09)	0.40(0.10)	0.54
Numbers of pauses	24.72(8.60)	23.83(8.64)	-0.94

Note: ***p < .001, **p < .01, *p < .05.

As for the results of gestures (Table 3), first, broadcasters produced a higher proportion of representational (23.55% vs. 17.72%, β=0.41, p<.001) and pointing gestures (19.58% vs. 17.67%, β=0.28, p<.001) in CDB than in ADB. Second, the average duration of each representational gesture in CDB was 180 ms longer than in ADB (p=.002). Additionally, gestures in CDB were more salient (p=.006) (see Figure 2). These results demonstrated that broadcasters relied more on representational gestures and produced more salient gestures to convey their information to children. By contrast, broadcasters produced a higher proportion of pragmatic gestures in ADB (37.11%) than in CDB (29.69%) (β=0.44, p<.001). The differences were not significant for the use of beat gestures (CDB: 25.1%, ADB: 27.5%, β=-0.10, p=.053). However, broadcasters reduced the number of repetitions within a beat gesture in CDB (p=.012).

Table 3: Means (SD) and results (beta and p) of gestures for each condition.

Dep variables	ADB	CDB	β and p
Repre_rate (per sec)	0.07(0.06)	0.10(0.08)	2.51***
Point_rate (per sec)	0.06(0.05)	0.08(0.07)	0.02***
Prag_rate (per sec)	0.19(0.11)	0.12(0.10)	-0.04***
Beat_rate (per sec)	0.11(0.06)	0.10(0.07)	-0.01
Average_beat_times	2.26(1.52)	1.94(0.89)	-0.32**
Mean duration of rep. (sec)	1.17(0.59)	1.35(0.54)	0.19**
Gesture saliency	6.32(0.97)	6.55(1.43)	0.23**

Note: ***p < .001, **p < .01, *p < .05.

Figure 2: A broadcaster explaining "two rabbits" in ADB (left) and CDB (right).

Discussion

This study provides a first insight into how broadcasters adapt their multimodal language to adults and in live broadcasting contexts. We found that broadcasters provided relatively less multimodal information when there was some assumed shared knowledge (i.e., addressing adult audiences). However, in the broadcasting context, broadcasters need to prioritize the completion of broadcasting content, thus their language production is not always audience-centred (e.g., speaking rate; pausing; lexical diversity) (Table 4).

Table 4: The choice of recipient design in multimodal cue

Multimodal Language	Features	Audience-centred
Linguistic features	MLU	√
	MATTR	ns
	Word frequency	√
	Question	√
Prosodic cues	Speaking rate	ns
	Pitch	√
	Intensity	ns
	pauses	ns
Gestures	Gesture type[1]	√
	Gesture rate[2]	√
	Beat gestures[3]	ns
	Repeat times in a beat	√
	Mean duration of rep.	√
	Gesture saliency	√

Note: ns = no significant evidence, [1]without including beat gestures; [2] without including beat gesture rate; [3] including the type of beat gestures and beat rate

Effects of common ground on recipient design

In line with previous research comparing CDL and ADL in non-broadcasting contexts, we discovered that broadcasters used shorter utterances, high(er)-frequency vocabulary (cf. Jaffe et al., 2001; Phillips, 1973; Snow, 1972) and more questions (cf., e.g., Newport, Gleitman, & Gleitman, 1977; Soderstrom et al., 2008), carried a higher pitch and more significant pitch variation (cf. Cotter, 1993; Grawunder et al., 2008; Medrado, Ferreira, & Behlau, 2005; Strangert, 2005), and produced more representational and pointing gestures (cf. Gogate, Bahrick, & Watson, 2000; Matatyaho-Bullaro et al., 2014; So et al., 2012) in CDB than in ADB. It has been claimed that CDS utterances tend to be shorter and grammatically simpler than those in ADS (Fletcher & MacWhinney, 1996; Saxton, 2009; Snow & Ferguson, 1977). When hosting a programme aimed at children, broadcasters used more common and frequent words as well as shorter utterances in CDB. Usually, children have a smaller vocabulary size than adults, which means that they may need help processing longer utterances and low-frequency words. Also, broadcasters must complete their speech in a limited time, so they may adopt shorter utterances and higher frequency words to make their speech easily understood by children while conveying as much information as possible.

There was a larger proportion of questions in CDB than in ADB. Questions used in CDS are a helpful type of input as they challenge children to reason and provide verbal

explanations (e.g., Duong et al., 2021; Purpura & Ganley, 2014; Tompkins et al., 2017; Rowe, Leech, & Cabrera, 2016). For example, caregivers' use of questions is shown to relate to children's unknown word learning (Dong et al., 2021). Although broadcasters cannot get audiences' feedback when hosting a programme, they still care about what children might think of and respond to their speech, as children may have a poorer comprehension of news than adults. Moreover, even if the primary goal of broadcasters' use of questions is not to get feedback or promote language acquisition, it can maintain or attract children's attention. Children have smaller mnemonic spans (Cowan et al., 1991) and constraints in working memory (Hulme & Tordoff, 1989; Luciana & Nelson, 1998), which allow them to hold less information compared to adults—asking a question benefits learning because it is an open communication form that allows children to generate hypotheses (Yu et al., 2018). To keep their attention, despite raising their pitch, broadcasters can also ask questions to promote thinking and keep children engaged in watching. Questions can be used as a language strategy when broadcasters organize their speech.

Intonation used in broadcasting discourse is usually carried by a higher pitch and a larger pitch variation than in ordinary conversations (Price, 2008). In the context of CDB, the pitch level is even higher, and the pitch variation is larger when broadcasters deliver children's programmes. However, we found an inconsistent finding with previous research. Researchers have argued that CDS has a wider pitch range than ADS (e.g., Fernald et al., 1989; Han et al., 2022; Narayan & McDermott, 2016), but in our study, CDB had a smaller pitch range than ADB. This is likely due to a ceiling effect that limited the broadcasters to enlarging the pitch maximum. The pitch maximum for CDB in our study was 0.47 semitone higher than ADB ($p<.001$), but the pitch minimum was 0.89 semitones higher than ADB ($p<.001$). The increase in pitch maximum in the CDB was not as significant as the pitch minimum, thus making the pitch range smaller when addressing children audiences.

Furthermore, broadcasters make use of gestures to help explain their speech. Campis and Özyürek (2013) found that the rate of iconic gestures increased for children when explaining how to use a coffee machine compared to when explaining to adults. Our study shows that broadcasters still produced more iconic and pointing gestures for the children's programme, even though no addressee was presented. In addition, broadcasters often exaggerate the crucial features when hosting children's programs. Besides raising pitch to attract their young audiences, they would also make gestures with larger sizes and extend the duration of representational gestures to make them clearer and more vivid to children. Also, pointing could map what they say with what is straightforwardly shown in the pictures to help children's understanding. The common ground between speakers and addressees affects various aspects of prosody and linguistic contents (see Isaacs & Clark, 1987) and gestures.

Effects of context requirement on recipient design

In contrast to the above results, we also found broadcasters did not adapt their programmes in some features between the two conditions. This suggests that the language strategies broadcasters adopted were not always designed especially for audiences, they still needed to take the unique parts of the broadcasting context into consideration.

In linguistic features, broadcasters did not show a significant difference in lexical diversity between CDB and ADB, whereas earlier research showed that CDS was significantly less diverse (Henning, Striano, & Lieven, 2005; Hills, 2012; Kaye, 1980; Phillips, 1973). The inconsistency may be explained by language modality and context. For example, in written CDS, picture books for children contain more unique words than oral speech (Cameron-Faulkner & Noble, 2013; Hayes & Ahrens, 1988; Montag, 2019; Montag, Jones, & Smith, 2015). In book-reading contexts, picture-book provides more lexically sophisticated speech than other caregiver-child activities (Crain-Thoreson, Dahlin, & Powell, 2001; Salo et al., 2015; Sosa, 2016; Weizman & Snow, 2001). It could be the case that, like picture books, broadcasting programmes contain prepared drafts before live broadcasting. Even in our experiment, participants had two minutes to prepare. Also, the degree of formality between broadcasting contexts and ordinary conversations directed to children is different. Speech used in broadcasting programs is usually well-organized and prepared.

In prosody, broadcasters did not slow down their speaking rate or reduce the number/ rate of pauses, and pause duration in CDB. CDS usually has a slower speaking rate and longer pauses in many languages compared to ADS (e.g., Bernstein-Ratner, 1985; Fernald & Simon, 1984; Marklund et al., 2014; Sjons et al., 2017; Tang & Maidment, 1996). One explanation for the differences between our and past results may be that, unlike stress-timed languages (such as English), Mandarin is closer to the typological extreme of syllable-timed languages (e.g., Mok & Dellwo, 2008). In stress-timed languages, stressed and unstressed syllables are distinguished in terms of duration and syllable weight, while syllable-timed languages have nearly equal weight and time in all syllables. Also, the durational variability is more remarkable in stress-timed languages compared to syllable-timed languages (Grabe & Low, 2002). For example, Han et al. (2022) showed that the speaking rate of CDS is not consistently slower than ADS across languages, ages, or across the board. They compared Dutch and Chinese and found that Dutch CDS was generally slower than ADS, while Chinese CDS did not. Another explanation may be that speech on television news often displays at a fast pace and is lack of pauses compared to non-announcing speech (e.g., Bolinger, 1982, 1989; Castro et al., 2010; Van Leeuwen; 1984). The nature of broadcasting thus does not allow much room for the adjustment of speaking rate or pausing, especially since the time setting was 1 minute for both programmes. The design of multimodal utterances has been examined in both speech and co-speech gestures. For example, people increase both speech and gesture efforts in noisy environments (i.e., an increase in speech intensity and

a pitch shift, see Trujillo et al., 2021; Rasenberg et al., 2022), and people use fewer words and gestures as common ground increases (Holler & Bavelas, 2017). However, unlike pitch modulation, speaking rate exhibits a trade-off between efficiency and information transmission. A slower speaking rate will sacrifice efficiency if a speaker is restricted to time. The speaker will tend to speak faster and reduce the pause duration to complete the speech under a given short-time instruction, while pitch adaptation will not take notably more time but only more effort.

Even though the proportions of beat gestures did not differ significantly in the two programmes, broadcasters beat more repeatedly and rapidly within a beat gesture in ADB. Beat gestures are usually used to bring listeners' attention to certain highlighted parts and emphasize information in speech (Ekman, 1999; McNeill, 1992; Wang & Chu, 2013). So et al. (2012) have found that beat gestures support memory, but beats seem to provide benefits to preschool children only when presented in pragmatically relevant contexts (Igualada, Esteve-Gibert, & Prieto, 2017). For native speakers and language learners, Rohrer, Delais-Roussarie, and Prieto (2020) found that beat gestures had no robust effect on information recall for native listeners and even led to significantly lower comprehension and memory recall for non-native language learners. When beat gestures are repeatedly used in discourse, they inherently lose their saliency as markers of important information (Rohrer et al., 2020). Due to the limited cognitive ability, children, to some degree, can be viewed as language learners analogous to adult second language learners who also experience cognitively more demands in an L2. Thus, too frequent beat gestures may increase the perceivers' cognitive load, which may hinder comprehension. As broadcasters, encoding words with appropriate beat gestures could help reinforce their content and, consequently, better emphasize the important content for their audience. However, broadcasters seem to avoid repeating beat gestures too much in CDB as it may decrease its power of marking information structure for children.

We found a novel result that broadcasters used more pragmatic gestures in ADB but tried to avoid such gestures in CDB. This may be because children do not use pragmatic gestures as proficiently as adults (e.g., Dowling, 2022). Unlike referential gestures such as iconic and deictics that serve cognitive benefits (e.g., memory, So et al., 2012) for audiences (Goldin-Meadow, 2003; Hadar & Butterworth, 1997; Kita, 2009), pragmatic gestures neither indicate or iconically represent a referent in conversation, nor do they have a conventionally agreed upon symbolic meaning (Ferré, 2012; Kendon, 2004, 2017; McNeill, 1992; Prieto et al., 2018; Shattuck-Hufnagel et al., 2016). In particular, children's ability to understand pragmatic gestures is poorer than adults' when a gesture's meaning is abstracted from its context. It has also been found that speakers produce pragmatic gestures, relating not to lexical content but to aspects of complicated interaction (Graziano & Gullberg, 2013). When talking to adults, broadcasters are talking to more competent audiences who share more common ground with them than children. They may have more difficulties speaking to competent audiences while speaking to children is more relaxed and effortless.

Why do broadcasters adapt some features for their audiences (audience-centred) but others not? As seen from the discrepancies in the above results, broadcasters were not constantly adapting to addressees. The distinguishing features of broadcasting may help answer the question. The primary goal of broadcasters is to produce their programmes of a good "quality". Therefore, they may give an equal weight to adults and children in some features of their productions. For example, the degree of diversity in their vocabulary and how they highlight prominent content in their programmes can also be viewed as factors contributing to good programmes. The trade-off effect on efficiency and effort can also play a role in adopting language strategies. Within the time limitation, sometimes broadcasters must sacrifice speaking rate or pause duration to complete their speech. Finishing a speech on time also contributes to maintaining the "quality" of live-broadcasting programmes. Besides, the notion of global adaptation, which is highly linked to the notion of expectation (Brennan, Galati, & Kuhlen, 2010), may also contribute to the complexity of recipient design. Global adaptation happens at the beginning of planning communicative expression and is grounded on speakers' expectations for different audience groups and individuals (Campis & Özyürek, 2013). In our experiment, even though broadcasters produced the same pictures for adults and children, they may expect that their child audiences have less background information about the pictures with them so they would produce CDB with more iconic gestures, rely more on higher-frequent words and shorter utterances. That is to say, recipient design is affected by the addressee and there is a direct relationship between different addressees and broadcasters' language choices. Furthermore, there are some features that the adaptation to the audience design was not significant. However, this does not mean that broadcasters do not engage in audience design or can automatically consider it as being speaker-centred. From a view of multimodal production in broadcasting language, it is driven at least in part by addressees, some of the signal channels are not audience-oriented whereas at the same time, other signal channels compensate for this through an audience design (e.g., gesture duration vs. speech rate). Thus, the distinction between speaker or audience orientation is not binary, but their mechanism should be understood as the incorporation of both processes simultaneously via different information channels according to contexts.

In sum, our study made the first attempt to document multimodal child-directed and adult-directed communication in a broadcasting context. It provides a better understanding of recipient design, which should be studied multimodally and viewed holistically across signal channels. Moreover, differences between CDS and CDB patterns could be attributed to the role of child-directed broadcasting in children's learning and development, and future studies can also expand to gender differences in multimodal broadcasting.

References

Abelman, R., & Atkin, D. (2000). What children watch when they watch TV: Putting theory into practice. *Journal of Broadcasting & Electronic Media*, *44*(1), 143-154.

Adams, C., Smith, M. C., Pasupathi, M., & Vitolo, L. (2002). Social context effects on story recall in older and younger women: Does the listener make a difference? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(1), 28-40.

Ambrazaitis, G., & House, D. (2017). Multimodal prominences: Exploring the patterning and usage of focal pitch accents, head beats and eyebrow beats in Swedish television news readings. *Speech Communication*, *95*, 100-113.

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, *92*(2), 484-501.

Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012). Audience design affects acoustic reduction via production facilitation. *Psychonomic Bulletin & Review*, *19*(3), 505-512.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31-56.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*, 1-22.

Bernstein-Ratner, N. (1985). Dissociations between vowel durations and formant frequency characteristics. *Journal of Speech, Language, and Hearing Research*, *28*(2), 255-264.

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

Bolinger, D. (1982). Intonation and its parts. *Language*, *58*(3), 505.

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford University Press.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, *53*, 301-344.

Castro, L., Serridge, B., Moraes, J., & Freitas, M. (2010). The prosody of the TV news speaking style in Brazilian Portuguese. In A. Botinis (Ed.), *Third ISCA Workshop on Experimental Linguistics* (pp. 17-20). University of Athens.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*(6), e10729.

Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and child directed speech. *First Language*, *33*(3), 268-279.

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.

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.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J. F. Leny & W. Kintsch (Eds.), *Language and comprehension* (pp. 287-296). New York, NY: Elsevier

Cotter, C. (1993). Prosodic aspects of broadcast news register. *Annual Meeting of the Berkeley Linguistics Society*, *19*(1), 90.

Covington, M. A., & McFall, J. D. (2010). Cutting the gordian knot: The moving-average type-token ratio (MATTR). J*ournal of Quantitative Linguistics*, *17*(2), 94-100.

Cowan, N., Saults, J. Scott., Winterowd, C., & Sherk, M. (1991). Enhancement of 4-year-old children's memory span for phonologically similar and dissimilar word lists. *Journal of Experimental Child Psychology*, *51*(1), 30-52.

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.

Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A. (2001). Parent-child interaction in three conversational contexts: Variations in style and strategy. *New Directions for Child and Adolescent Development*, *2001*(92), 23-38.

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

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*(43).

Dowling, N. (2022). *Obviously I don't know but whatever: Emblematic and pragmatic uses of shrug gestures in early childhood and adolescence*. Doctoral dissertation, Department of comparative human development, University of Chicago.

Duong, S., Bachman, H. J., Votruba-Drzal, E., & Libertus, M. E. (2021). What's in a question? Parents' question use in dyadic interactions and the relation to preschool-aged children's math abilities. *Journal of Experimental Child Psychology*, *211*, 105213.

Ekman, P. (1999). Emotional and conversational nonverbal signals. In L. S. Messing & R. Campbell (Eds.), *Gesture, speech, and sign* (pp. 45–55). Oxford, UK: Oxford University Press.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, *20*(1), 104-113.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477-501.

Ferré, G. (2012). Functions of three open-palm hand gestures. *Journal Multimodal Communication*, *1*, 5-20.

Fletcher, P., & Macwhinney, B. (1996). *The handbook of child language*. Oxford: Blackwell.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35-51.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Gogate, L. J., Bahrick, L. E., & Watson, J. D. (2000). A study of multimodal motherese: The role of temporal synchrony between verbal labels and gestures. *Child Development*, *71*(4), 878-894.

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby) talk to me. *Current Directions in Psychological Science*, *24*(5), 339-344.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *laboratory phonology* (pp. 515-546). Berlin: Mouton de Gruyter.

Grawunder, S., Engert, H., Kaiser, H., Unger, S., & Bose, I. (2008). Pitch, speech rate and rhythm as speaking style characterizing parameters in German news presentation. [Poster presentation]. *EASR08*, *UCL*, London.

Graziano, M. (2014). The development of two pragmatic gestures of the so-called open hand Supine family in Italian children. In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance*: *Essays in honor of Adam Kendon* (pp. 311-330). John Benjamins Publishing Company.

Graziano, M., & Gullberg, M. (2013). Gesture production and speech fluency in competent speakers and language learners. [Paper presentation]. *Tilburg Gesture Research Meeting (TiGeR) 2013*, Tilburg.

Gupta, R. K., Saini, D. P., Acharya, U., & Miglani, N. (1994). Impact of television on children. *Indian Journal of Pediatrics*, *61*(2), 153-159.

Hadar, U., & Butterwork, B. (1997). Iconic gestures, imagery, and word retrieval in speech. *Semiotica*, *115*, 147-172.

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

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.

Hayes, D. P., & Ahrens, M. G. (1988). Vocabulary simplification for children: a special case of "motherese"?. *Journal of Child Language*, *15*(2), 395-410.

Henning, A., Striano, T., & Lieven, E. V. M. (2005). Maternal speech to infants at 1 and 3 months of age. *Infant Behavior and Development*, *28*(4), 519-536.

Hills, T. (2012). The company that words keep: comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language*, *40*(3), 586-604.

Hoff, E., & Krueger, W. M. (1991). Older siblings as conversational partners. *Merrill-Palmer Quarterly*, *37*(3), 465-481.

Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, *26*(1), 4-27.

Holler, J., & Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, *24*(2), 267-289.

Holler, J., & Bavelas, J. (2017). Multi-modal communication of common ground: A review of social functions. In R. B. Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 213-240). Amsterdam: Benjamins.

Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length, and acoustic similarity on serial recall. *Journal of Experimental Child Psychology*, *47*(1), 72-87.

Igualada, A., Esteve-Gibert, N., & Prieto, P. (2017). Beat gestures improve word recall in 3- to 5-year-old children. *Journal of Experimental Child Psychology*, *156*, 99-112.

Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, *116*, 26-37.

Izci, B., Jones, I., Özdemir, T. B., Alktebi, L., & Bakir, E. (2019). Youtube and young children: Research, concerns and new directions. *Crianças, famílias e tecnologias. Que desafios? Que caminhos?*, 81-92.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), 1-132.

Jenvey, V. B. (2007). The relationship between television viewing and obesity in young children: a review of existing explanations. *Early Child Development and Care*, *177*(8), 809-820.

Kaye, K. (1980). Why we don't talk 'baby talk' to babies. *Journal of Child Language*, *7*(3), 489-507.

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. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

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

Kita, S. (2009). Cross-cultural variation of speech-accompanying gesture: A review. *Language and Cognitive Processes*, *24*(2), 145-167.

Kiepura, E., Niedźwiecka, A., & Kmita, G. (2021). Silence matters: The role of pauses during dyadic maternal and paternal vocal interactions with preterm and full-term infants. *Journal of Child Language*, *49*(3), 1-18.

Lamb, M. E., & Lewis, C. (2010). The development and significance of father-child relationships in two-parent families. In M. E. Lamb (Ed.), *The role of the father in child development* (pp. 94-153). John Wiley & Sons Inc.

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

Luciana, M., & Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, *36*(3), 273-293.

Marklund, U., Marklund, E., Lacerda, F., & Schwarz, I. (2014). Pause and utterance duration in child-directed speech in relation to child vocabulary size. *Journal of Child Language*, *42*(5), 1158-1171.

Masataka, N. (1992). Motherese in a signed language. *Infant Behavior and Development*, *15*(4), 453-460.

Matatyaho-Bullaro, D. J., Gogate, L., Mason, Z., Cadavid, S., & Abdel-Mottaleb, M. (2014). Type of object motion facilitates word mapping by preverbal infants. *Journal of Experimental Child Psychology*, *118*, 27-40.

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

McNeill, D. (1998). Speech and gesture integration. In J.M. Iverson, S. Goldin-Meadow, et al. (Eds.), *The nature and functions of gesture in children's communication* (pp. 11-27). San Francisco: JosseyBass.

Medrado, R., Ferreira, L. P., & Behlau, M. (2005). Voice-over: Perceptual and acoustic analysis of vocal features. *Journal of Voice*, *19*(3), 340-349.

Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, *39*(5), 527-546.

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear. *Psychological Science*, *26*(9), 1489-1496.

Morris, D., Collett, P., Marsh, P., & O'Shaughnessy, M. (1979). *Gestures, their origins and distribution*. New York: Stein and Day.

Mok, P.P., & Dellwo, V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures: Cantonese, beijing Mandarin and English. In P. Barbosa, S. Madureira, & C. Reis (Eds.), *Proceedings of the 4th conference on speech prosody* (pp. 423-426). Editoria RG/CNPq.

Mok, P.P., Fung, H.S., Li, J. (2014) A preliminary study on the prosody of broadcast news in Hong Kong Cantonese. In N. Campbell, D. Gibbon & D. Hirst (Eds.), *Proceedings of the 7th conference on speech prosody* (pp. 1072-1075). Trinity College Dublin.

Müller, C. (1998). Iconicity and gesture. In S. Santi, I. Guatiella, C. Cave, & G. Konopczyncki (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication. Actes du colloque [Orality and gestuality: Multimodal interaction and behavior in communication. Proceedings of the meeting of ORAGE 2001]* (pp. 407–410). Paris, France: L'Harmattan.

Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. The *Journal of the Acoustical Society of America*, *139*(3), 1272-1281.

Neumann, M. M., & Herodotou, C. (2020). Evaluating YouTube videos for young children. *Education and Information Technologies*, 25.

Newport, E. L., Gleitman, H., & Gleitman, L. R. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 109-149). Cambridge, England: Cambridge University Press.

Perniss, P., Lu, J. C., Morgan, G., & Vigliocco, G. (2017). Mapping language to the world: the role of iconicity in the sign language input. *Developmental Science*, *21*(2), e12551.

Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development*, *44*(1), 182.

Price, J. (2008). New news old news: A sociophonetic study of spoken Australian English in news broadcast speech. *AAA: Arbeiten Aus Anglistik Und Amerikanistik*, *33*(2), 285-310.

Prieto, P., Cravotta, A., Kushch, O., Rohrer, P., & Vilà-Giménez, I. (2018). Deconstructing beat gestures: A labelling proposal. *Proceedings 9th International Conference on Speech Prosody* (pp. 201–205). Poznan, Poland: ISCA.

Purpura, D. J., & Ganley, C. M. (2014). Working memory and language: Skill-specific or domain-general relations to mathematics? *Journal of Experimental Child Psychology*, *122*, 104-121.

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

Raneri, D., Von Holzen, K., Newman, R., & Bernstein Ratner, N. (2020). Change in maternal speech rate to preverbal infants over the first two years of life. *Journal of Child Language*, *47*(6), 1263-1275.

Rasenberg, M., Pouw, W., Özyürek, A., & Dingemanse, M. (2022). The multimodal nature of communicative efficiency in social interaction. *Scientific Reports*, *12*(1), 19111.

Rodero, E., & Cores-Sarría, L. (2021). Best prosody for news: A psychophysiological study comparing a broadcast to a narrative speaking style. *Communication Research*, *0*(0).

Rohrer, P. L., Delais-Roussarie, E., & Prieto, P. (2020). Beat gestures for comprehension and recall: Differential effects of language learners and native listeners. *Frontiers in Psychology*, *1*,1.

Rowe, M. L., Leech, K. A., & Cabrera, N. (2016). Going beyond input quantity: Wh-questions matter for toddlers' language and cognitive development. *Cognitive Science*, *41*, 162-179.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696-735.

Salo, V. C., Rowe, M. L., Leech, K. A., & Careda, N. J. (2015). Low-income fathers' speech to toddlers during book reading versus toy play. *Journal of Child Language*, *43*(6), 1385-1399.

Saxton, M. (2009). The inevitability of child directed speech. *Language Acquisition*, 62-86.

Schubotz, L., Özyürek, A., & Holler, J. (2018). Age-related differences in multimodal recipient design: younger, but not older adults, adapt speech and co-speech gestures to common ground. *Language, Cognition and Neuroscience*, *34*(2), 254–271.

Schubotz, L., Özyürek, A., & Holler, J. (2022). Individual differences in working memory and semantic fluency predict younger and older adults' multimodal recipient design in an interactive spatial task. *Acta Psychologica*, *229*, 103690.

Shattuck-Hufnagel, S., Ren, A., Mathew, M., Yuen, I., & Demuth, K. (2016). Non-referential gestures in adult and child speech: Are they prosodic? In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Proceedings from the 8th international conference on speech prosody* (pp. 836–839). Boston: Boston University.

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

Shute, B., & Wheldall, K. (1999). Fundamental frequency and temporal modifications in the speech of British fathers to their children. *Educational Psychology*, *19*(2), 221-233.

Shute, B., & Wheldall, K. (2001). How do grandmothers speak to their grandchildren? Fundamental frequency and temporal modifications in the speech of British grandmothers to their grandchildren. *Educational Psychology*, *21*(4), 493-503.

Sim, J. H. (2021). Sociophonetic variation in English /l/ in the child-directed speech of English-Malay bilinguals. *Journal of Phonetics*, *88*, 101084.

Sjons, J., Hörberg, T., Östling, R., & Bjerva, J. (2017). Articulation rate in Swedish child-directed speech increases as a function of the age of the child even when surprisal is controlled for. In M. Włodarczak (Ed.), *Proceedings of the 18th annual conference of the international speech communication association* (pp. 1794-1798). Red Hook, NY: Curran Associates, Inc.

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, *43*(2), 549.

Snow, C. E., & Ferguson, C. A. (1977). *Talking to children: language input and acquisition*. Cambridge University Press.

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., Blossom, M., Foygel, R., & Morgan, J. L. (2008). Acoustical cues and grammatical units in speech to two preverbal infants. *Journal of Child Language*, *35*(4), 869-902.

Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *The Journal of the Acoustical Society of America*, *128*(1), 389-400.

Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent-infant communication. *JAMA Pediatrics*, *170*(2), 132.

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.

Strangert, E. (2005). Prosody in public speech: analyses of a news announcement and a political interview. *Proceedings of the 6th Annual Conference of the International Speech Communication Association* (pp. 3401-3404). Lisbon, PT: International Speech Communication Association.

Swerts, M., & Krahmer, E. (2010). Visual prosody of newsreaders: Effects of information structure, emotional content and intended audience on facial expressions. *Journal of Phonetics*, *38*(2), 197-206.

Swerts, M., & Marsi, E. (2012). Prosodic evaluation of accent distributions in spoken news bulletins of Flemish newsreaders. *The Journal of the Acoustical Society of America*, *132*(4), 2616-2624.

Tang, J. S. Y., & Maidment, J. A. (1996). Prosodic aspects of child-directed speech in Cantonese. *University College London: Speech, Hearing and language—Work in Progress*, *9*, 257-276.

Tippenhauer, N., Fourakis, E. R., Watson, D. G., & Lew-Williams, C. (2020). The scope of audience design in child-directed speech: Parents' tailoring of word lengths for adult versus child listeners. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(11), 2163-2178.

Tompkins, V., Bengochea, A., Nicol, S., & Justice, L. M. (2017). Maternal inferential input and children's language skills. *Reading Research Quarterly*, *52*(4), 397-416.

Trujillo, J., Özyürek, A., Holler, J., & Drijvers, L. (2021). Speakers exhibit a multimodal Lombard effect in noise. *Scientific Reports*, *11*(1).

Van Der Geest, T. (1977). Some interactional aspects of language acquisition. In C. Snow & C. Ferguson, (Eds.), *Talking to Children: Language Input and Acquisition*. Cambridge: Cambridge University Press.

Van Leeuwen, T. (1984). Impartial speech: Observations on the intonation of radio newsreader. *Australian Journal Cultural Studies*, *2*(1), 84-98.

Van de Mieroop, D., E. Zenner, & S. Marzo. (2016). Standard and Colloquial Belgian Dutch pronouns of address: A variationist-interactional study of child-directed speech in dinner table interactions. *Folia Linguistica*, *50*(1), 31-64.

Wang, L., & Chu, M. (2013). The role of beat gesture and pitch accent in semantic processing: An ERP study. *Neuropsychologia*, *51*(13), 2847-2855.

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology*, *37*(2), 265-279.

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* (LREC 2006) (pp. 1556-1559).

Yu, Y., Landrum, A. R., Bonawitz, E., & Shafto, P. (2018). Questioning supports effective transmission of knowledge and increased exploratory learning in pre-kindergarten children. *Developmental Science*, *21*(6), e12696.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*(2), 179-208.