

Article

Multimodal Dimensions of Hungarian Infant-Directed Communication During Storytelling

Anna Kohári ^{1,*}, Uwe D. Reichel ¹, Sarolta Murányi ², Rebeka Dóra Keserű ², Katalin Pirsell ^{1,2}
and Katalin Mády ¹

¹ ELTE Research Centre for Linguistics, H-1068 Budapest, Hungary; uwe.reichel@nytud.elte.hu (U.D.R.); pirsellk@gmail.com (K.P.); mady.katalin@nytud.elte.hu (K.M.)

² Department of Applied Linguistics and Phonetics, ELTE Eötvös Loránd University, H-1088 Budapest, Hungary; muranyi.sarolta@btk.elte.hu (S.M.); keseru.rebeka343@gmail.com (R.D.K.)

* Correspondence: kohari.anna@nytud.elte.hu

Abstract

Adults, when talking to infants, adjust their communication across several dimensions, including acoustic, linguistic, and gestural characteristics. However, our knowledge of how these dimensions covary in this register adjustment remains limited. Our objective was to analyze the linguistic elements, gestures and acoustic features (speech rate, vowel space, prosody) of storytelling directed to 6-month-old infants in a multidimensional way. The storytelling of 92 mothers to their own children was investigated and compared to their communication with other adults. We analyzed the acoustic features, gestures, and the linguistic construction of their storytelling. Our results have shown that pointing and showing gestures were more common in infant-directed storytelling, just as interactive expressions unrelated to the story were. The correlation tests across dimensions revealed that infant-directed communication is multidimensional. The cross-dimensional analyses demonstrated that infant-directed communication, compared to adult-directed communication, exhibited more systematic multidimensional coordination. Mothers who used more gestures and interactive elements in their infant-directed communication also tended to modify the acoustic characteristics of their speech to a greater extent relative to their adult-directed speech.

Keywords: infant-directed speech; acoustic analysis; prosody; gesture; multimodality; Hungarian



Academic Editors: Jing Shao, Susan Rvachew and Margaret Kehoe

Received: 28 February 2026

Revised: 28 June 2026

Accepted: 22 July 2026

Published: 4 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Speakers tend to adjust the characteristics of their communication depending on the situation and on the perceived or actual needs of their counterparts, in order to communicate successfully and to express social closeness or distance (Gallois & Giles, 2015; Pickering & Garrod, 2006). For example, adults talk differently to infants than to other adults. This specialized speech register is commonly referred to as infant-directed speech (IDS). Caregiver–infant interactions are typically multidimensional, as adults modify not only the acoustic properties of their speech, but also their gestures and explicit linguistic expressions. Nevertheless, a substantial body of existing research has focused primarily on a single dimension, namely, acoustic characteristics (Cox et al., 2023b; Fernald et al., 1989; Hilton et al., 2022; Miyazawa et al., 2017). Relatively few studies have examined the multidimensional nature of infant-directed communication (IDC) using an integrative

approach (e.g., [de La Cruz-Pavía et al., 2020](#); [Smith & Strader, 2014](#)). Therefore, we seek to gain a better understanding of how the different dimensions of communication directed at infants are interrelated, and what strategies speakers employ across these dimensions to ensure successful communication.

1.1. Properties of Infant-Directed Communication

Speech acoustics is one of the most thoroughly studied features of IDC. Results have shown that in the languages studied (e.g., English, Spanish, and Polish) IDS is typically slower than adult-directed speech (ADS), and exhibits generally higher fundamental frequency ([Cox et al., 2023a](#); [Hilton et al., 2022](#); [Kuhl et al., 1997](#); [Raneri et al., 2020](#)). However, findings on whether the fundamental frequency (f0) range is different in the two registers are inconsistent. In some studies, speakers' f0 range in IDS was higher than in ADS (e.g., in American English, Mandarin Chinese, Polish, and Norwegian; [Hilton et al., 2022](#); [Rosslund et al., 2024](#)), while other studies investigating the infant-directed speech of Cantonese-, Kenyan-, Fijian- and American English-speaking mothers found no such relationship ([Broesch & Bryant, 2015](#); [Xu & Burnham, 2010](#)). The results for vowel production are variable across studies. The vowel space, calculated based on the distances between vowel formants, was found to be larger in IDS compared to ADS in American and Australian English, Japanese, Russian, and Swedish ([Cristia & Seidl, 2014](#); [Kalashnikova & Burnham, 2018](#); [Kuhl et al., 1997](#); [Miyazawa et al., 2017](#)). However, other studies did not find such an effect (e.g., in Norwegian: [Englund & Behne, 2005](#); in English, French and Japanese: [Dodane & Al-Tamimi, 2007](#)). Although the acoustic properties of IDS change as the child ages, in the first year, the difference from ADS is manifested generally in the features mentioned above ([Cox et al., 2023a](#)). Although only a very limited number of systematic comparisons exist between the acoustic features of different infant-directed communication situations, the available results suggest that infant-directed speech shows highly similar characteristics in free play and storytelling ([Gergely et al., 2017](#); [Rosslund et al., 2024](#)). Several characteristics discussed here were observed across languages ([Fernald et al., 1989](#); [Hilton et al., 2022](#)), although substantial language-specific or cultural differences may also exist ([Broesch & Bryant, 2015](#); [Narayan & McDermott, 2016](#)). Therefore, we address the characteristics of infant-directed speech produced by native Hungarian speakers specifically. In Hungarian infant-directed speech, as in other languages, lower speech rate and higher fundamental frequency were reported ([Gergely et al., 2017](#); [Kohári et al., 2024](#)). Hungarian infant-directed speech also exhibits a larger vowel space compared to adult-directed speech ([Gergely et al., 2017](#); [Kohári et al., 2024](#)). These features of infant-directed speech appeared in spontaneous speech situations and reading tasks, in which participants were asked to read sentences aloud ([Gergely et al., 2017](#)). In terms of f0 range, however, studies found conflicting results for Hungarian speech. While one study reported a larger f0 range in Hungarian native speakers' infant-directed speech than in their adult-directed speech ([Gergely et al., 2017](#)), a study using a different method for assessing f0 range (measuring f0 in semitones to account for inter-speaker variability) found no such difference ([Kohári et al., 2024](#)).

Infant-directed communication, however, is not restricted to the acoustic properties of speech. Adults also tend to adjust their nonverbal signals and gestures when communicating with infants ([Iverson et al., 1999](#); [Murphy, 1978](#); [Rohlfing et al., 2015](#); [Schmidt, 1996](#); [Vigliocco et al., 2019](#)). Motion-tracking analyses have revealed that parents performed actions with objects for longer durations and at closer distances when communicating with their infants than when interacting with adults ([van Schaik et al., 2020](#)). Based on their referential values, gestures are classified into four types: deictic, iconic, metaphoric, and beat gestures ([McNeill, 1992](#)). The two main subtypes of deictic (pointing and showing) gestures occur frequently in IDS; furthermore, iconic and conventional gestures can also

be found (Iverson et al., 1999; Rohlfing et al., 2015; Vigliocco et al., 2019). In contrast, other studies have reported such a low frequency of deictic gestures in IDC that they were excluded from analysis, and only iconic and beat gestures were investigated (Kandemir et al., 2024). These discrepancies may arise as the usage of gestures varies considerably depending on the communication situation (e.g., story reading, storytelling, and free play), as well as on the child's age (Iverson et al., 1999; Kandemir et al., 2024; Rohlfing et al., 2015). Different situations elicited different gesture patterns. It has been found that in the book-reading condition, mothers used pointing gestures much more frequently than in free play (Rohlfing et al., 2015). Another study showed that during story reading and storytelling, the number of gestures produced by speakers—particularly iconic gestures—was markedly lower than when they were asked to explain how to navigate to a given location (Kandemir et al., 2024). The emergence and the frequency of different gesture types may be largely influenced by whether the object is present or far away from the speaker. Very few studies have investigated deictic gestures—and pointing gestures in particular—in IDC during the early stage of language development, before the infant reaches one year of age (Deák et al., 2018; Murphy, 1978; Schmidt, 1996). This gap in the literature is especially surprising given that infants themselves begin to produce pointing gestures as early as nine months of age (Kita, 2003), presumably after learning them from caregiver input. This raises the question of what we actually know about deictic gestures in speech directed to infants aged nine months or younger. In one study examining communication with infants aged 3 to 11 months during free play in a fixed seating arrangement, caregivers produced pointing gestures rather rarely. The authors suggested that this may have been due to the recording procedure and the physical distance between the mother, the infant, and the objects present in the setting (Deák et al., 2018). In contrast, another study found that during joint book reading, caregivers already produced pointing gestures in speech directed to nine-month-old infants, and that the frequency of pointing increased substantially by 24 months of age (Murphy, 1978). Similarly, in a study investigating communication with preverbal 10-month-old infants during free play, gestures were found to be more frequent than pointing gestures (Schmidt, 1996). The results of this study also suggested a developmental increase in the use of pointing gestures over time (Schmidt, 1996). Since the use of deictic gestures during storytelling may vary substantially across languages and linguistic communities (Nicoladis et al., 2018; Pika et al., 2006), it is important to note that, to the best of our knowledge, gestures in Hungarian IDC have not yet been investigated.

Infant-directed communication may differ not only in its acoustic and gestural characteristics but also in its linguistic expressions. Linguistic research may encompass a wide range of approaches; for instance, studies have shown that adults typically simplify syntactic structures in child-directed speech and frequently repeat words and expressions when addressing children (Genovese et al., 2020; Harmati-Pap et al., 2025; Saint-Georges et al., 2013; Soderstrom, 2007). Storytelling and joint picture-book reading, however, represent highly specific linguistic contexts characterized by a fixed structure and, in many cases, printed text on the pages (McArthur et al., 2005; Wigglesworth & Stavans, 2001). Narrative construction involves certain key elements that are indispensable for comprehending the story. At the same time, storytelling may also include utterances whose communicative function is not the transmission of information but the gaining and maintaining of attention (Wigglesworth & Stavans, 2001). These storytelling and interactive storybook-reading strategies have typically been investigated among preschool-aged children (Stavans, 1996; Stavans & Goldzweig, 2008; Tompkins et al., 2017). Their findings indicate that parents of children aged three years and older predominantly produce story-related utterances, although they also incorporate interactive elements intended to actively involve the child in communication (Stavans & Goldzweig, 2008). The communicative differences between

interactions with older children and those with preverbal infants are particularly evident in the fact that, during storytelling with older children, adults are able to ask genuine questions and expect verbal responses (Kang et al., 2009). A substantial proportion of caregivers reported that they had begun shared picture-book reading before the child had reached six months of age and that this activity had become part of their everyday routines (Tompkins et al., 2017). A study specifically addressed conversational-function utterances in infant-directed storytelling that explicitly contained lexical items such as “look” or “see” (Murphy, 1978). The authors found that such utterances occurred more frequently in speech directed to nine-month-old infants than at later ages (14, 20, and 24 months). In speech directed to preverbal infants, these forms have been described as salient features based on qualitative analyses (Goldring-Zukow, 1991).

1.2. Interactional Functions of Infant-Directed Communication

Traditionally, three main functions are associated with the characteristics of IDC, namely attentional, affective, and didactic purposes, which are not mutually exclusive (Kempe et al., 2024; Saint-Georges et al., 2013; Spinelli et al., 2017). The different characteristic features can serve similar or identical functions simultaneously in communication, supplementing each other. In what follows, we briefly review the earlier findings and their conclusions regarding the possible functions of the acoustic, gestural, and linguistic features of IDC.

In infant-directed communication, adults spontaneously adjust the features of their speech to the situation (Hilton et al., 2022). In interaction, even infants below one year of age are capable of providing feedback, for example through gaze or head-turning. It has been observed across languages that infants prefer infant-directed speech over adult-directed speech, and that this preference increases with development between 4 and 14 months of age (ManyBabies Consortium, 2020). Even newborns are able to distinguish the typical acoustic properties of infant-directed speech from those of adult-directed speech (Håden et al., 2020). However, it is possible that not all characteristics are equally preferred by infants. It has been observed, using manipulated rather than spontaneous speech, that 6-month-old infants listened longer to slowed-down infant-directed speech with artificially exaggerated pitch contour, whereas 10-month-olds preferred a normal tempo and more monotonous speech (Kitamura & Notley, 2009). Furthermore, it has been found that 6–8-month-old infants prefer vowels with infant-like formants, meaning formants that are more similar to their own, over typical adult formants (Polka et al., 2022). Beyond the characteristics of infant-directed speech, infants also showed preferences in infant-directed actions over the features of adult-directed communication. In a study, infants were shown video recordings of an adult manipulating an object under infant-directed (ID) and adult-directed (AD) conditions. Both groups of 6–8-month-old and 11–14-month-old infants preferred ID actions over AD actions (Brand & Shallcross, 2008). Besides ID actions, infants at 4–5 months of age are capable of shifting their visual attention in the direction of a pointing gesture (Rohlfing et al., 2012). Not only gestures, but also the characteristics of infant-directed speech can direct the child’s attention to an object. Six-month-old infants tended to follow an adult’s shift in gaze toward an object more after hearing a word produced with the acoustic features of infant-directed speech than when it was pronounced in the manner of adult-directed speech (Senju & Csibra, 2008). Although specific phrases were not investigated, expressions with appropriate meaning (e.g., ‘Look baby, look!’ in Rohlfing et al., 2012) may also help directing the infants’ gaze, and for this reason such phrases were used as a part of the experiment protocol. The characteristics of IDC can thus signal that the child is the addressee of the communication, and some of these features may even serve to direct the infant’s attention toward objects.

In IDC, the speaker may also convey emotions. Although infant-directed speech has very similar characteristics to those associated with the expression of positive emotions, in IDC expressing strong emotions is not necessarily the case. One study attempted to separate the emotional and non-emotional components of infant-directed speech (Parlato-Oliveira et al., 2020). Infants under 1 year of age showed a listening preference for emotional speech over neutral speech, mostly when prosodic patterns associated with happiness or sadness were applied; in these cases, the infants exhibited longer listening time (Kao et al., 2022). The emotion-signaling function of IDC gestures, however, has not yet been investigated extensively in a similar manner, to the best of our knowledge.

Regarding the features of infant-directed communication, it has also been suggested that they may be related to the infant's later vocabulary. It is to be emphasized, however, that the correlations found between the characteristics of infant-directed speech and infant's expressive or receptive vocabulary do not necessarily imply a causal connection or any corresponding intention on the part of the speaker (cf. Kempe et al., 2024). Acoustic analyses have shown that slower speech rate, higher fundamental frequency, and larger vowel space are correlated with the child's later receptive and expressive vocabulary (Hartman et al., 2017; Mengru & Kager, 2024; Raneri et al., 2020; Tóth et al., 2025). Besides the acoustic properties, pointing gestures directed toward an object or event also were associated with the child's vocabulary (Rohlfing et al., 2015; Salo et al., 2019). The properties of IDC may thus play similar roles in communication. However, a question arises as to whether, and to what extent, these properties appear simultaneously in communication and whether their effects may enhance or complement each other.

1.3. Multidimensional Analyses of Infant-Directed Speech

IDC exhibits a wide range of distinctive features, including acoustic properties, gesture characteristics, and specific explicit linguistic expressions. It is well established that during accommodation to a given situation or interaction partner, speakers may modify multiple communicative features (acoustic, nonverbal, etc.) simultaneously. It is not necessary that speakers express the same interactional function across multiple dimensions simultaneously, but convergence indicating a speaker's cooperation can be considered the default form of accommodation. This cooperative attitude can appear simultaneously in numerous speech characteristics (speech rate, pitch, vowel formants, etc.) during interaction (Pellegrino et al., 2024). This also holds for infant-directed speech. Besides the acoustic features, gestures also exhibit observable differences in infant-directed communication compared to adult-directed communication (Iverson et al., 1999; Rohlfing et al., 2015; Vigliocco et al., 2019). Gestures and prosody form an organic whole in terms of various functions, complementing and enhancing each other—often simultaneously—in communication between adults (Prieto et al., 2025), and the same close connection can also be assumed in infant-directed communication. Furthermore, gestures often accompany the pronunciation of words not only in adult-directed communication, but also in infant-directed communication (e.g., pointing at and naming an object) (Kosie & Lew-Williams, 2024; Murphy, 1978). Prosody cannot be separated from words; although we investigate them separately, changes in prosodic features are linked to prominent words or phrase boundaries (Prieto et al., 2025). Moreover, the acoustic, nonverbal, and linguistic features of IDC serve similar communicative functions. The features that adults adjust in infant-directed communication (e.g., speech rate, pitch, vowel space, gestures) tend to be listened to longer, and preferred by the infants over the features of adult-directed speech (Kitamura & Notley, 2009; Polka et al., 2022; Rohlfing et al., 2012). They may even give feedback about their preferences to the speaker, for example through gaze or head-turning, which can further enhance the use of these communicative tools.

In the light of the connections discussed above, it is somewhat surprising that only rather sparse data exist on the strategies speakers apply when adapting different communicative characteristics to the situational context and to the infant addressee. Only a small number of pioneering studies have examined the relationships between multiple dimensions of IDC to one another, and to the acoustic properties of speech in particular. In one study investigating speech directed to eight-month-old infants, head movements and fundamental frequency were found to be more strongly related to each other in IDC than in adult-directed communication (ADC) (Smith & Strader, 2014). In another study, participants were asked to address a child shown on a screen in an imagined scenario, and the researchers examined head nods and movements, as well as eyebrow movements. They found that acoustically marked prominence was not accompanied by such gestures, whereas utterance-final positions were marked both acoustically and by head nods (de La Cruz-Pavía et al., 2020). Connections between explicit linguistic forms and gestures have also been reported. In communication directed to infants between 9 and 24 months of age, it has been shown that lexical items explicitly containing *look* or *see* are frequently accompanied by pointing gestures (Murphy, 1978). A comprehensive investigation of nonverbal features in speech addressed to 18–24-month-old children showed that approximately 60% of speech co-occurred with some form of nonverbal cue (Kosie & Lew-Williams, 2024). To summarize, the co-occurrence of features across different communicative dimensions appears to be common in IDS. However, relatively little is known about how acoustic properties relate to other communicative features, particularly in the case of preverbal infants.

1.4. The Current Study

The primary objective of the present study was to gain a better understanding of the adaptation characteristics of speakers communicating with their own children, and to explore how they modify the features of the different communicational dimensions. Infant-directed communication can happen in a wide variety of contexts and activities. During early language development, one common everyday situation is shared picture-book viewing and storytelling based on picture books (Tompkins et al., 2017). This context provides an excellent opportunity to investigate communicative characteristics within a relatively well-controlled setting, with comparable textual content and contextual constraints across participants.

Infants are able to distinguish between speech addressed to them and speech addressed to adults, and, as indicated by their attention span, they typically prefer the typical characteristics of infant-directed communication over those of adult-directed communication (Kitamura & Notley, 2009; Polka et al., 2022; Rohlfing et al., 2012). Thus, any differences between the two registers, or the presence of typical features of infant-directed communication, may represent additional opportunities to grab and maintain the attention of the infants. However, relatively little is known about the gestural and linguistic characteristics of storytelling directed to preverbal infants. Thus, it was also our objective to analyze these understudied features during the early stages of language development. These communication features cannot be considered independent of each other. Instead, they occur together during the adaptation to infant-directed communication, serving similar communicational purposes and complementing each other. Redundant information, even when conveyed through different channels, may facilitate speech processing and information transfer (cf. Alviar et al., 2026). Therefore, we analyzed the connections between gestures, linguistic characteristics of storytelling, and acoustic features in order to gain a deeper understanding of the multidimensional nature of infant-directed speech. Speakers may adopt different strategies when accommodating to their interaction partner: they may modify the characteristics of their speech and nonverbal communication to a varying extent, depending on

the perceived or actual needs of the addressee. Speakers tend to spontaneously emphasize certain parts of their speech to direct attention, and this marking of prominence may occur through prosody (including timing), gestures (Prieto et al., 2025), or even explicit linguistic phrases. Therefore, among other questions, we sought to determine whether speakers who use attention-directing nonverbal gestures and explicit linguistic elements are also more likely to use typical prosodic properties of infant-directed speech. Alternatively, it is possible that some speakers primarily rely on explicit linguistic forms to attract attention, while others put greater emphasis on nonverbal cues or prefer to exploit prosodic features.

2. Materials and Methods

2.1. Participants

We analyzed audio and video recordings of 92 first-time mothers. All speakers were monolingual Hungarian speakers without any diagnosed hearing, speech, or other language-related disorders. The participants were residents of Budapest or its metropolitan area and ranged in age from 22 to 44 years (mean (M) = 31.3 years, standard deviation (SD) = 4.9). Recruitment took place at the Birth Centre of the Military Hospital in Budapest, Hungary at the time of the birth of their first child. None of the infants was born preterm. The recordings were conducted when the infants were 6 months old (M = 6 months and 22.8 days, SD = 6.8, 40 male). This age was chosen because infants at this stage of language development are already able to recognize words but do not yet actively produce them (Bergelson & Swingley, 2012, 2015; Tincoff & Jusczyk, 2012). Thus, at this age, infants are already interested in the words they hear in a storytelling situation, but this may still be a highly active phase of language development, as the segmentation of new and unfamiliar words may also be required. Furthermore, such pre-lexical infants are capable of shifting their visual attention toward an object in response to a pointing gesture (Rohlfing et al., 2012); therefore, adults communicating with them may already receive some feedback—even at this early age—about whether the children follow the interaction to some extent. It has been observed that infants prefer the characteristics of infant-directed speech over adult-directed speech, although this preference may change over time. In the selected age range, however, a clear preference of the characteristics of infant-directed speech can be assumed based on the results of previous studies (Kitamura & Notley, 2009).

2.2. Materials and Procedure

We collected audio and video recordings in a laboratory environment when the infants were 6 months old. First, the mothers were given a storybook containing 16 pictures, 10 of which were accompanied by caption texts (Figure 1). The storybook was created by the project members with the assistance of a professional illustrator, in accordance with the objectives of our investigation (see details in the Supplementary Materials). The mothers were asked to tell a whole coherent story in their own words (producing spontaneous speech) while incorporating the written captions. These mandatory sentences ensured identical phonetic environments for the acoustic measurements across all speakers and conditions. The pictures ensured controlled narrative content and cognitive load, while the spontaneous segments yielded a more natural speaking style. After becoming thoroughly familiar with the book, the mothers told the story first to another adult, then to their child, holding the book in their hands. The pictures and the read sentences ensured comparable story content in both situations, enabling us to compare the properties of adult-directed (AD) and infant-directed (ID) communications. The differences between these two registers indicate whether, and to what extent, the speakers adjusted their communicative behavior.



Figure 1. The pages of the storybook. The translation of the Hungarian sentence on the above storybook page is as follows: ‘The three pixies ran away from the lion as fast as they could.’

For the audio recordings, we used a head-mounted hypercardioid microphone (Beyer-dynamic TG H74c), and the speech was recorded using Audacity at a 44.1 kHz sampling rate and 16-bit resolution via an USB external sound card (M-Audio, two-channel). In both situations, the mothers sat facing their speech partner, and the infants were seated in a high-chair, in accordance with one established practice in infant-directed communication research (Abu-Zhaya et al., 2017). The video recordings captured a side view of the participants via a CX550 camera mounted on a fixed tripod, enabling us to track the mother’s hand movements (Figure 2).



Figure 2. A camera shoot of an adult–adult communication.

2.3. Data Processing and Analysis

2.3.1. Acoustic Analysis of Infant-Directed and Adult-Directed Speech

The acoustic properties were analyzed based on the read sentences using Praat (version 6.2.17; (Boersma & Weenink, 2019)). For each speaker and each recorded story, the boundaries and the text of 10 read sentences (containing 17 clauses) were manually annotated and automatically segmented using the WebMAUS service version 3.11 (Kisler et al., 2017). Based on previous research (Fernald et al., 1989), speech rate, vowel space, and the median, maximum, minimum, and range of f_0 may differ between adult-directed and infant-directed speech; therefore, these measures were calculated). Vowel formant and fundamental frequency data were extracted using Praat. F_0 values were converted to semitones to control for inter-speaker variability. The phonetic parameters were calculated using the CoPaSul toolkit version 1.3.0 (Reichel, 2016), which handles f_0 -related measurement errors (e.g., octave errors and outliers) and provides robust f_0 estimates. Speech rate (in syllables per second) was defined as the number of syllables in a read sentence divided by the total duration of the utterance. The extracted f_0 -related measures and speech rate values were averaged separately for the two registers. Vowel space was calculated based on the F_1

and F2 values across the entire story told by the mothers. We computed the mean Euclidean distance of the F1–F2 values from the F1–F2 centroid (similarly to Bradlow et al., 1996).

2.3.2. Analysis of Infant-Directed Gestures

The type, starting point and end point of the mothers' gestures were annotated in ELAN version 6.8; see Lausberg and Sloetjes (2009). During storytelling, the mother was required to keep the picture book in hand, which considerably constrained their gestures; nevertheless, they were able to point with their index fingers. The onset and offset of these pointing gestures were annotated. Here, a gesture is defined as a sequence of movements between two rest positions (Kendon, 1980). When a speaker did not produce a pointing gesture, two possible scenarios were distinguished. In the first case, the mother held the book pages in a way that made them visible to the speech partner (Gupta et al., 2025). This was labeled as a special type of showing gesture in which the pages of the book are displayed (cf. Karmazyn-Raz & Smith, 2023). When neither pointing nor showing gestures were produced, the section was assigned a "no gesture" label. Thus, the three categories (pointing gesture, showing gesture, and no gesture) were mutually exclusive. Two independent annotators analyzed the data from 25 speakers, and their gesture classifications showed high inter-rater agreement (Cohen's kappa: $\kappa = 0.91$). The average number of words in storytelling was 306 ($SD = 127$) in the AD condition and 357 ($SD = 134$) in the ID condition. The average duration of the storytelling task was 175.0 s ($SD = 72.5$) in the ADC and 212.0 s ($SD = 88.0$) in the IDC. The mean number of pointing gestures produced during storytelling was 0.4 ($SD = 1.7$) in the ADC and 7.5 ($SD = 7.8$) in the IDC. The mean number of showing gestures produced during storytelling was 0.4 ($SD = 1.9$) in the ADC and 8.3 ($SD = 7.9$) in the IDC. The pointing rate was calculated for each speaker in each register by dividing the number of pointing gestures produced in a given story by the total number of words that each individual caregiver produced when they told that story. Analogously, the showing rate was calculated based on the frequency of showing gestures. We also determined the percentage of the duration of (pointing and showing) gestures relative to the total time of the storytelling.

2.3.3. Narrative Analysis of Infant-Directed Speech

In our investigation, we used storytelling as a context, which is a common everyday activity for infants (Tompkins et al., 2017) and provides an excellent opportunity to examine communicative characteristics in a relatively well-controlled setting, with comparable textual content and contextual constraints across participants. However, this setting also created a unique linguistic situation, as storytelling has a distinctive structure with certain fixed elements (McArthur et al., 2005; Wigglesworth & Stavans, 2001). To exclude non-essential elements of the story that may have attention-grabbing meanings or interactional functions, we analyzed its underlying structure using the tools and methods of story grammar.

For the narrative analysis of the stories told by the mothers, we used the manually corrected transcripts of the recorded tales. The narratives were segmented into so-called communication units, following common practice in infant-directed speech research (Cebioğlu et al., 2022; Loban, 1976). Hereafter, we refer to these communication units as utterances, a term widely used in phonetics. Coordinate clauses were treated as separate utterances, whereas subordinate clauses were considered part of the same utterance.

In total, the speakers produced 11,079 utterances. The mean number of utterances per story produced by the speakers was 52.5 ($SD = 15.6$) in the AD condition and 68.0 ($SD = 26.1$) in the ID condition. Based on the pictures in the storybook, we identified the narrative elements essential for telling and understanding the story (McArthur et al.,

2005). Narrative elements that did not appear in the read sentences, and were therefore present only in spontaneous speech (e.g., walnut, hat, ship, bird, lion, etc.), were identified. In addition, we also marked those utterances whose content was not strictly connected to the storytelling but instead served to enhance the infant's engagement (Wigglesworth & Stavans, 2001). These were classified as conversational utterances (e.g., *Can you see that? Look, here are the pixies! Here we go, let us turn the page!*). The read sentences did not contain any conversational elements. As the identification of conversational utterances may involve subjective judgment, two of the authors independently annotated 33.3% of the data. Interrater agreement was high (Cohan's kappa: $\kappa = 0.99$). For each story produced by a speaker, we calculated the number of utterances containing narrative elements and the number of conversational utterances. We then determined the proportion of narrative and conversational utterances within each story, taking the total number of utterances in that story as 100%.

2.3.4. Statistical Analyses

For all three types of measures (acoustic, gestural, and storytelling-related), we assessed the associations between properties of adult-directed and infant-directed communication within each speaker. As a large proportion of these measures was not normally distributed, and because we aimed to compare effect sizes, paired Wilcoxon signed-rank tests were used in all cases in R version 4.5.2 (R Core Team, 2025). For the paired Wilcoxon tests, the appropriate effect size tests were performed using the *rcompanion* package version 2.5.2 (Mangiafico, 2016). If the value of r was found larger than 0.60, the effect size was considered large. Values between 0.40 and 0.60 were considered medium, and values between 0.10 and 0.40 were classified as small (Mangiafico, 2016). To examine the association between the different dimensions, we first calculated the difference between the two registers ($ID - AD$) for each measure. These differences were then analyzed using Spearman's rank correlation, and we organized the resulting values in a correlation matrix using the *corrplot* package version 0.95 (Wei & Simko, 2024). p -values were adjusted using FDR (False Discovery Rate) correction to account for multiple comparisons. Spearman correlation coefficient indicates a strong relationship when its magnitude is over 0.7, a moderate correlation when it is between 0.30 and 0.70, and weak below 0.30 (Akoglu, 2018).

3. Results

3.1. Acoustic Differences Between Infant- and Adult-Directed Speech

First, similarly to previous studies, we investigated whether the acoustic features of infant-directed speech differ from those of adult-directed speech. Furthermore, we analyzed how consistently speakers differentiated between the two registers. We calculated the average of the acoustic measurements separately for speakers in the two registers. Table 1 shows the mean values and standard deviations obtained for all speakers in the two registers; therefore, the standard deviations in the table represent between-speaker variability. However, the results of the statistical tests are derived from paired comparisons conducted within speakers to explore the extent of consistency in the speakers' differentiation of the registers. In agreement with earlier findings (e.g., Cox et al., 2023a; Fernald et al., 1989; Kalashnikova & Burnham, 2018), almost all acoustic features in the present dataset showed significant inter-register differences (Table 1). Speech rate was lower in infant-directed speech, whereas vowel space area was larger than in adult-directed speech. The minimum, maximum and median values of f_0 were all higher in infant-directed speech than in the other register; the f_0 range, however, did not show such a difference. Based on the large within-speaker effect sizes obtained for speech rate, vowel space area, and the minimum, maximum, and median values of f_0 , these differences appeared to be rather systematic (Table 1).

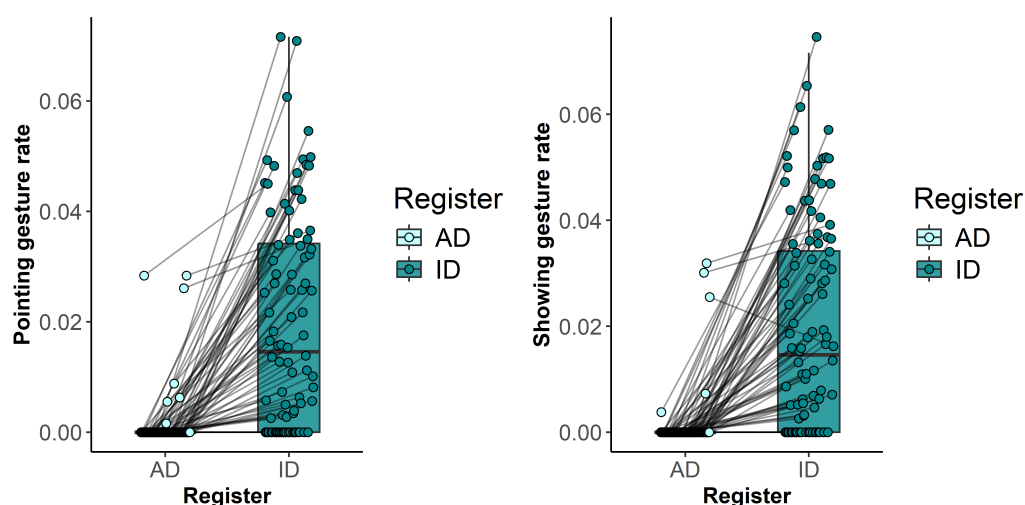
Table 1. Mean and standard deviation (in parentheses) of acoustic measures in IDS and ADS. All measures were computed separately for each speaker. Statistical test results are provided in the table.

Measure	IDS <i>M</i> (<i>SD</i>)	ADS <i>M</i> (<i>SD</i>)	Statistic Test Result	Effect Size (<i>r</i>)
Speech rate	3.8 (0.5)	4.2 (0.5)	$V = 192$	−0.788 ***
Vowel space	2.1 (0.2)	2.0 (0.2)	$V = 3837$	0.723 ***
F0 median	12.7 (3.9)	11.2 (3.8)	$V = 3932$	0.763 ***
F0 minimum	4.6 (3.4)	3.0 (3.3)	$V = 3874$	0.638 ***
F0 maximum	22.7 (4.2)	21.8 (4.4)	$V = 3350$	0.669 ***
F0 range	5.3 (1.2)	5.6 (1.5)	$V = 1732$	n.s.

*** $p < 0.001$, n.s. = non-significant

3.2. Gestural Differences Between Infant- and Adult-Directed Communication

In the mothers' communication, we also compared whether the frequency of the different gesture types (pointing gesture rate, showing rate) and the gesture duration proportions relative to the entire story exhibit differences between adult-directed and infant-directed communication. Speakers typically produced more pointing gestures in infant-directed communication than in adult-directed communication (IDC: $M = 0.020$, $SD = 0.019$; ADC: $M = 0.001$, $SD = 0.005$). However, it should be noted that 22 speakers (23.9%) did not produce any of the investigated gestures in infant-directed communication (Figure 3). An additional five speakers (5.4%) only showed the pages of the book but did not produce pointing gestures. The showing gesture rate was generally higher in IDC than in ADC (IDC: $M = 0.022$, $SD = 0.020$; ADC: $M = 0.001$, $SD = 0.005$). Both measures differed significantly in the two registers based on the statistical tests (pointing gesture rate: ($V(92) = 2145$; $p < 0.001$, $r = 0.869$); showing gesture rate: ($V(92) = 2469$; $p < 0.001$, $r = 0.858$)). Thus, the majority of the speakers systematically produced pointing or showing gestures more often in IDC than in ADC. Besides the frequency of the gesture occurrences, the combined total duration of the gestures relative to the duration of the story was also analyzed. We calculated gesture (pointing and showing gesture) duration as a percentage of total text duration (%) in a story (Figure 4). This measure was also found to be higher in IDC than in ADC (IDC: $M = 69.1$, $SD = 42.7$; ADC: $M = 2.4$, $SD = 13.7$). As confirmed by the statistical analyses, speakers consistently produced gestures for a longer proportion of the storytelling time in IDC than in ADC ($V(92) = 2485$; $p < 0.001$, $r = 0.873$).

**Figure 3.** Occurrence frequency of pointing and showing gestures relative to the total number of words in the story in infant-directed and adult-directed speech. Each point represents one speaker's value in a given register, and the lines connect the two registers for the same speaker.

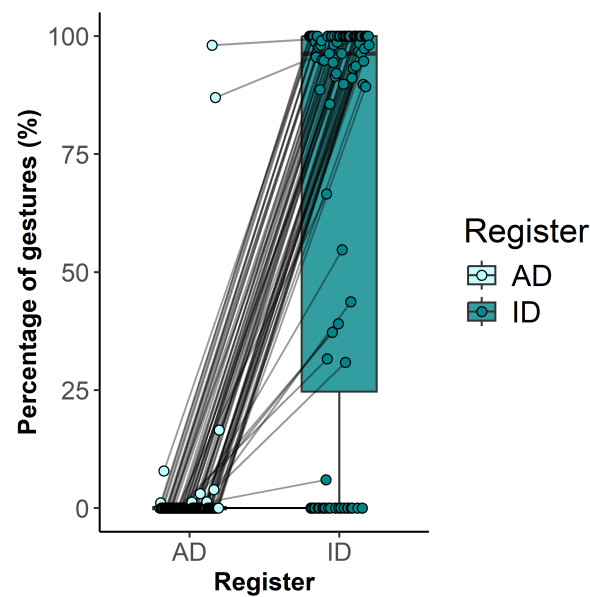


Figure 4. Gesture (pointing and showing gesture) duration as a percentage of total text duration (%) in a story. Each point represents one speaker's value in a given register, and the lines connect the two registers for the same speaker.

3.3. Narrative Differences Between Infant- and Adult-Directed Speech

We also examined whether the two registers differed in their use of narrative and conversational elements. We evaluated the percentage of utterances within a given story that contained narrative elements. Narrative elements were defined as key units of information for the storytelling. In our analysis we considered only those elements that were not among the written sentences provided in the storybook, as we focused on the characteristics of spontaneous speech production. On average, in a speaker's infant-directed speech, 10.2% of the utterances ($SD: 6.4$) contained at least one of the specified narrative elements, and the same proportion was observed in adult-directed speech ($M = 10.2\%$, $SD: 5.5$). The frequency of narrative elements did not differ significantly between the two registers ($V(92) = 1771$; $p = 0.55$).

Besides the storytelling-related elements, other linguistic expressions also appeared in the narratives that served to involve the infant in the storytelling. The parent intends to attract and maintain the child's attention with these linguistically explicit expressions. We therefore calculated the percentage of the utterances that contained such conversational elements or phrases. On average, infant-directed speech was found to contain a higher proportion of conversational elements ($M = 9.7\%$, $SD = 10.7$), than adult-directed speech ($M = 1.4\%$, $SD = 2.8$). This feature showed significant differences between the two registers in the statistical analyses ($V(92) = 2898$; $p < 0.001$, $r = 0.808$). Most speakers generally used few conversational elements when telling the story to an adult, whereas they used more in infant-directed storytelling (Figure 5). Although the group mean showed deviations between the two registers, it is important to note that 16 speakers (17.4%) did not produce any conversational elements in infant-directed speech either. Among those who did, the proportion was consistently higher in infant-directed than in adult-directed speech.

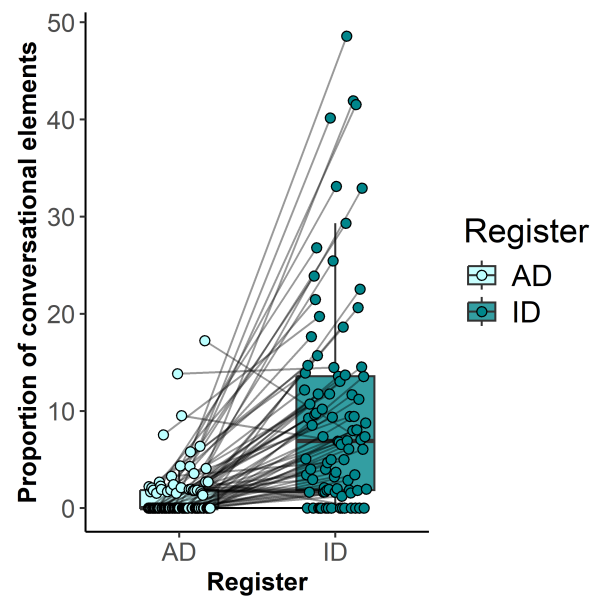


Figure 5. Proportion of conversational elements in infant-directed and adult-directed speech. Each point represents one speaker's value in a given register, and the lines connect the two registers for the same speaker.

3.4. Assessing the Relationship Between the Different Dimensions of Infant-Directed Speech

Infant-directed communication differs from adult-directed communication in several aspects. This raises the question of whether the different dimensions of communication are related to each other. To explore this, we examined possible connections among the investigated properties. We approached the calculation of correlations from two perspectives. First, we evaluated the extent to which speakers differentiate between the two registers. For this, we considered inter-register differences ($ID - AD$) across all measures and calculated their correlations (Figure 6). However, we aimed not only to quantify how much speakers differentiate between the two registers but also to obtain typical within-speaker values resulting from adaptation to the infant-directed communication (IDC) situation and to explore the relationships between them (Figure 7). However, analyses conducted within one register may also require a reference condition; thus, we determined the values for the same speakers separately in the AD condition as well (Figure 8). Therefore, the data shown in Figures 7 and 8 are closely related and should be interpreted together.

First, we determined the correlations among the inter-register ($ID - AD$) differences for all measures (Figure 6). The results showed that the dimensions of infant-directed communication (acoustic, nonverbal, and storytelling-related measures) were correlated with each other in certain cases (Figure 6). For example, pointing and showing gesture rate and gesture proportion exhibited positive correlations. Similarly, the measures related to f_0 were found to be moderately intercorrelated based on the statistical tests. Within the group of acoustic measures, speech rate showed a negative correlation with both f_0 median and maximum. In other words, the more a speaker slowed down their speech in IDS relative to ADS, the more likely that speaker was to differentiate the two registers not only in vowel space but also in terms of f_0 median and maximum. The storytelling-related measures were similarly interrelated. The greater the proportion of conversational elements used by a speaker, the more likely it was that the proportion of narrative elements tended to decrease in infant-directed speech compared to adult-directed speech.

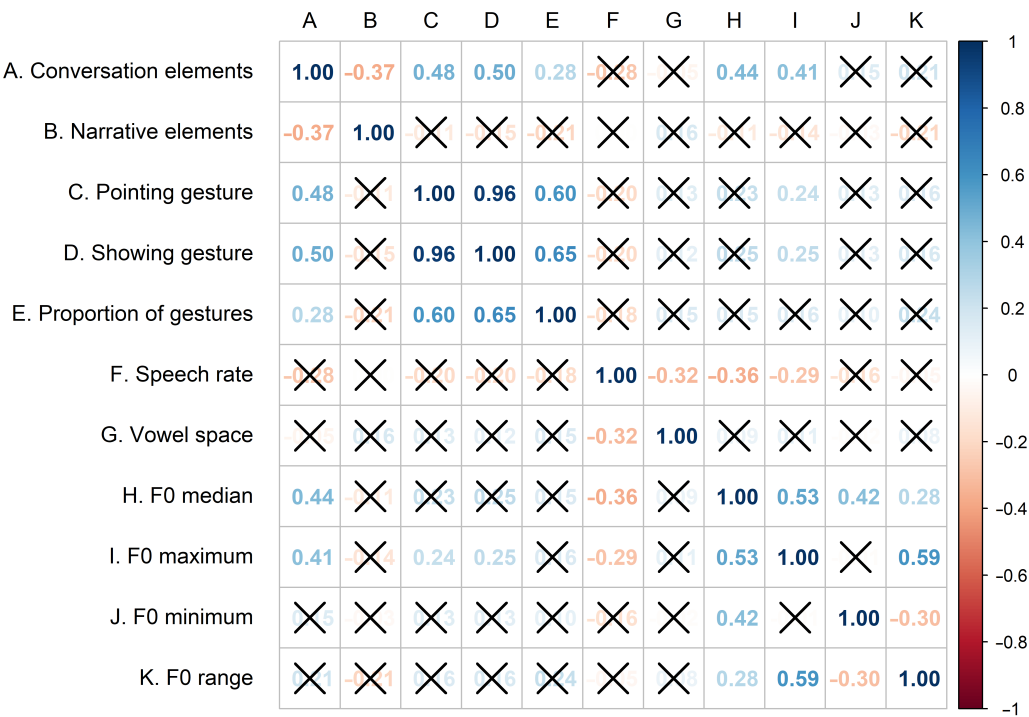


Figure 6. Results of correlation tests among measures calculated from the differences between the two registers. The values in the table indicate Spearman’s correlation coefficients. X denotes non-significant results.

The relationships between the characteristic features of infant-directed communication extend beyond the boundaries of a given dimension. A higher frequency of conversational elements was positively correlated with both gesture-related measures and two acoustic properties, namely f0 median and maximum. Thus, speakers who tend to use more conversational elements when telling stories to their children relative to their adult-directed speech also tend to gesture more frequently and for a greater proportion of time in IDC than in ADC. Moreover, these speakers are more likely to differentiate between the two registers in terms of f0-related acoustic properties, thereby using acoustic cues to attract and maintain the infant’s attention. Thus, speakers who distinguish between the two registers along one communicative dimension tend to exploit multiple dimensions of communication simultaneously.

We also examined the relationships among the measures of infant-directed communication to explore the correlations among the characteristic values associated with adaptation to the communicative situation. For reference, we performed the same correlation analyses for adult-directed communication. The correlations among the infant-directed measures were found to be similar in several respects to the pattern that can be obtained analyzing the differences between the two registers (Figure 7). Measures within the same communicative dimension showed intercorrelations. One notable difference, however, was that narrative elements and conversational elements did not correlate when only infant-directed speech was considered; thus, within a single register, they can be regarded as independent measures. The analysis of infant-directed communication further revealed that gestures were more closely related to acoustic features (e.g., vowel space, f0 minimum). The more frequently a speaker produced gestures, and the greater the proportion of time spent gesturing, the more likely they were to use an expanded vowel space and to exhibit an increased f0 minimum in infant-directed speech. Additional weak correlations are given in Figure 7.

Furthermore, the results shown in Figure 8 indicate that the measures in adult-directed communication were correlated with each other only within the same dimensions (acoustic, nonverbal, or storytelling-related).



Figure 7. Results of correlation tests among measures in IDC. The values in the table indicate Spearman’s correlation coefficients. X denotes non-significant results.



Figure 8. Results of correlation tests among measures in ADC. The values in the table indicate Spearman’s correlation coefficients. X denotes non-significant results.

4. Discussion

The main goal of the present study was to better understand the connections between the characteristic features of the speakers' adaptation when communicating with their children. Specifically, we examined the acoustic features, deictic gestures, and linguistic constructions of storytelling in caregivers' communication with their 6-month-old preverbal infants. We investigated the speakers' multidimensional communication strategies of accommodation. However, before conducting our multidimensional analyses, we evaluated which features are indeed present in infant-directed communication and differ from those observed in adult-directed communication.

In the present study, we explored separately which verbal, nonverbal, and acoustic properties were more specific to storytelling addressing 6-month-old infants than to adult-directed communication in Hungarian. It should be noted that the comparisons of the two conditions relied on telling the same fairy tale. Therefore, our findings may not extend to other communicative situations. In addition, since the storytelling from a book is unusual in adult-to-adult communication, the findings in ADC can be viewed merely as a reference point for assessing the characteristics of infant-directed communication. Considering these limitations, we obtained the following results. The acoustic differences found between the two registers were in agreement with results reported in the literature (Fernald et al., 1989; Hilton et al., 2022). Infant-directed speech was slower, exhibited a larger vowel space, and several f_0 -related properties (median, maximum, minimum) were also higher. However, the f_0 range did not show such inter-register differences (similarly to, e.g., Broesch & Bryant, 2015). Our data indicate that one possible explanation may be that, as both the minimum and maximum of f_0 is higher in IDS than in ADS, the range only gets shifted but does not change in size between the two registers. Besides the acoustic properties, we also investigated the deictic gestures (pointing, showing). Both the pointing and showing gestures appear more frequently and for a longer duration in storytelling directed to preverbal infants than in adult-directed communication. The frequent pointing gestures may provide appropriate input for infants, who by 9 months of age, typically start to actively use these gestures, presumably following the caregivers' example (Kita, 2003). Besides gestures, we also analyzed the stories from a linguistic perspective. The fundamental pieces of information required for storytelling were found in similar proportions in both registers; however, utterances facilitating interaction and unrelated to the story (e.g., *Can you see that? Look, here are the pixies! Here we go, let us turn the page!*) were more frequent in IDS than in ADS. Previous research has also observed that utterances containing the words *look* and *see* are more frequent in joint book reading with 9-month-old infants, and later, as the children age, the proportion of such utterances gradually decreases (Murphy, 1978). In the present study, however, we did not focus exclusively on specific words. Instead, from the perspective of story grammar, we annotated all optional utterances serving an interactive function, which differed markedly in frequency between the two registers. Although this feature is not typically highlighted among the canonical characteristics of IDS, our findings indicate that adults apply linguistically explicit, functionally interactive elements to facilitate the infant's engagement in communication. Future research could provide further quantitative pragmatic analyses in order to further explore the use and types of these linguistic devices.

Infant-directed communication can thus be considered multimodal, and our results also revealed that its specific features in different dimensions are interconnected. Redundant pieces of information serving the same function—even if they are transmitted through different channels—can aid the infant in perceiving the properties of communication (Bahrick et al., 2019). Therefore, we examined the association between the gestures, the linguistic features of storytelling, and its acoustic properties with one another in order to

advance our understanding of the multidimensionality of infant-directed communication. Within each individual dimension (gestural, acoustic, and narrative characteristics), the features were, unsurprisingly, interrelated. Among the acoustic features, the speech rate of IDS was found to be associated with vowel space, even though these measures are not trivially linked by definition. Slower IDS production typically coincided with an expanded vowel space compared to ADS. Reaching the articulatory targets of a vowel takes time, thus in faster speech, articulatory targets are often not completely realized. Similarly, if the intonation extends to higher fundamental frequencies, achieving these f_0 targets may likewise require additional time.

Our results indicate that infant-directed communication should not primarily be understood as a broad multidimensional coordination of features. Rather, it appears to be a system in which specific features from different dimensions are interconnected. Speakers who adapt their communication more strongly to the interactional context do not necessarily employ all available communicative resources simultaneously. The relationships between the characteristics of infant-directed communication extend across dimensions. The higher frequency of conversational elements showed positive correlations both with gesture-related measures and with two of the acoustic features, namely median and maximum f_0 . Thus, speakers who tended to use more conversational elements during storytelling directed to infants, relative to their adult-directed speech, also gestured more frequently and for a greater proportion of time when addressing infants than when addressing adults. In addition, these speakers were more likely to differentiate the two registers in terms of f_0 -related properties. Accordingly, speakers who differentiate between the two registers in terms of f_0 -related dimensions also tend to exploit multiple communicative dimensions simultaneously, supporting attention direction not only through prosodic cues but also through explicit linguistic expressions and gestures. Previous studies have demonstrated that speakers tend to spontaneously emphasize certain parts of their speech in adult-to-adult communication, and this marking of prominence may occur through prosody (including timing), gestures (Prieto et al., 2025), or even explicit linguistic phrases. More targeted investigation is required to determine whether our findings on multidimensional coordination can be explained by the prominence of word occurrences and their marking. When only IDS was considered, the frequency and relative duration of gestures during storytelling were associated with several acoustic features. The more frequently a speaker gestured, and the greater the proportion of time spent gesturing, the more likely they were to use an expanded vowel space and elevated f_0 -related values in infant-directed speech. It has been observed that infants prefer speech with the same acoustic and nonverbal properties (vowel space, f_0 -related features, gestures) over typical adult-directed communication (Brand & Shallcross, 2008; Kitamura & Notley, 2009; Polka et al., 2022). Therefore, when these characteristic features appear in speech—regardless of the extent to which it differs from adult-directed speech—infants in the interaction may even provide feedback about their preferences to the speaker, for example through gaze or head-turning, which can further enhance the use of these communicative tools. The relevance of these cross-dimensional correlations in IDC is underlined by the fact that in adult-directed communication, we observed significant correlations only within individual dimensions, but not across them.

Our findings on the correlations between characteristics of IDC also warrant caution when investigating the effects of any particular feature of infant-directed speech on language development, as several characteristics of infant-directed communication were found to exhibit correlations within and, in some cases, across dimensions. Consequently, even if one or a few features of infant-directed communication were shown to correlate with children's later vocabulary size, such correlations alone would not justify the con-

clusion that these features are causally related to subsequent vocabulary development. Although correlation analyses comparing characteristics of infant-directed speech with later vocabulary development undoubtedly help identify potential connections and provide valuable insights into communicative processes (Raneri et al., 2020; Rohlfing et al., 2015; Salo et al., 2019), their findings should primarily be regarded as starting points for further research. Rather than relying solely on indirect evidence, future work should aim to identify the mechanisms through which specific communicative features directly affect children's perception, attention, memory, etc. (cf. Kempe et al., 2024).

To summarize, our results have shown that during infant-directed communication adults adjust not only the acoustic features of their speech, but also the frequency and relative duration of their gestures, and this register also differs in terms of linguistic expressions from adult-directed communication. Although these dimensions are usually separate from each other, certain characteristics of the IDC of an adult have a marked tendency to co-occur with other features. A small number of the parents, however, did not include gestures when talking to their children at all, and typically they were the ones who used fewer conversational elements and differentiated less between IDC and ADC in terms of acoustic properties. Speakers thus generally follow different strategies when communicating with their children, and these communications strategies often appear together in dimensions with modality (in acoustic, linguistic, and nonverbal features).

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/languages11080163/s1>, Supplementary Material S1: The list of read sentences in the storybook in Hungarian; Supplementary Material S2: Gesture coding; Supplementary Material S3: Annotation of narrative elements.

Author Contributions: Conceptualization, A.K. and K.M.; methodology, A.K., S.M., U.D.R. and K.M.; software, U.D.R.; validation, A.K., K.P., R.D.K. and S.M.; formal analysis, A.K. and U.D.R.; investigation, A.K., S.M., R.D.K. and K.P.; resources, K.M. and U.D.R.; data curation, A.K., S.M., R.D.K. and K.P.; writing—original draft preparation, A.K., R.D.K. and K.P.; writing—review and editing, A.K., U.D.R. and K.M.; visualization, A.K.; supervision, A.K.; project administration, A.K.; funding acquisition, A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Hungarian National Research, Development and Innovation Office, grant number PD134775 and K115385.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Medical Research Council–Committee of Scientific and Research Ethics in Hungary (ETT-TUKEB approval number 17187-3) and the United Ethical Review Committee for Research in Psychology in Hungary (EPKEB approval number 77_2015).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Informed parental consent was obtained for the child data used.

Data Availability Statement: The data are not publicly available due to ethical considerations. Recordings and videos cannot be shared due to privacy reasons. The toolkit used for extracting phonetic parameters is available in the public domain: <https://github.com/reichelu/copasul> (accessed on 21 July 2026). The data generated during the analyses are available upon request from the authors.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

AD	adult-directed
ADC	adult-directed communication
ADS	adult-directed speech
F0	fundamental frequency
ID	infant-directed
IDC	infant-directed communication
IDS	infant-directed speech
M	mean
n.s.	non-significant
SD	standard deviation

References

- Abu-Zhaya, R., Seidl, A., & Cristia, A. (2017). Multimodal infant-directed communication: How caregivers combine tactile and linguistic cues. *Journal of Child Language*, 44(5), 1088–1116. [PubMed]
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91–93. [CrossRef] [PubMed]
- Alviar, C., Jones, W., & Lense, M. (2026). InCHORRRuS: Infant-directed communication highlights and organizes repetition and redundancy through rhythmic structure. *Annals of the New York Academy of Sciences*, 1556(1), e70147. [PubMed]
- Bahrack, L. E., McNew, M. E., Pruden, S. M., & Castellanos, I. (2019). Intersensory redundancy promotes infant detection of prosody in infant-directed speech. *Journal of Experimental Child Psychology*, 183, 295–309. [CrossRef] [PubMed]
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258. [CrossRef]
- Bergelson, E., & Swingle, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, 11(4), 369–380. [PubMed]
- Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* [Computer software manual]. Available online: <http://www.praat.org/> (accessed on 22 August 2022).
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, 20(3–4), 255–272. [CrossRef] [PubMed]
- Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science*, 11(6), 853–861. [CrossRef] [PubMed]
- Broesch, T. L., & Bryant, G. A. (2015). Prosody in infant-directed speech is similar across western and traditional cultures. *Journal of Cognition and Development*, 16(1), 31–43.
- Cebioğlu, 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), e622–e638. [CrossRef] [PubMed]
- Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., & Fusaroli, R. (2023a). A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour*, 7(1), 114–133.
- Cox, C., Diderksen, C., Keren-Portnoy, T., Roepstorff, A., Christiansen, M. H., & Fusaroli, R. (2023b). Infant-directed speech does not always involve exaggerated vowel distinctions: Evidence from Danish. *Child Development*, 94(6), 1672–1696. [CrossRef]
- Cristia, A., & Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *Journal of Child Language*, 41(4), 913–934. [CrossRef] [PubMed]
- Deák, G. O., Krasno, A. M., Jasso, H., & Triesch, J. (2018). What leads to shared attention? Maternal cues and infant responses during object play. *Infancy*, 23(1), 4–28.
- de La Cruz-Pavía, I., Gervain, J., Vatikiotis-Bateson, E., & Werker, J. F. (2020). Coverbal speech gestures signal phrase boundaries: A production study of Japanese and English infant-and adult-directed speech. *Language Acquisition*, 27(2), 160–186.
- Dodane, C., & Al-Tamimi, J. (2007). An acoustic comparison of vowel systems in adult-directed-speech and child-directed-speech: Evidence from French, English & Japanese. In *Proceedings of the 16th international congress of phonetics sciences* (pp. 6–10). Saarland University.
- Englund, K. T., & Behne, D. M. (2005). Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research*, 34(3), 259–280. [CrossRef] [PubMed]
- 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. [CrossRef] [PubMed]
- Gallois, C., & Giles, H. (2015). Communication accommodation theory. In *The international encyclopedia of language and social interaction* (pp. 1–18). John Wiley & Sons, Ltd.

- Genovese, G., Spinelli, M., Lauro, L. J. R., Aureli, T., Castelletti, G., & Fasolo, M. (2020). Infant-directed speech as a simplified but not simple register: A longitudinal study of lexical and syntactic features. *Journal of Child Language*, 47(1), 22–44. [PubMed]
- Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific Reports*, 7(1), 13739. [CrossRef] [PubMed]
- Goldring-Zukow, P. (1991). A socio-perceptual/ecological approach to lexical development: Affordances of the communicative context. *Anales de Psicología*, 7(2), 151–162.
- Gupta, S., Karnovska, S., Jover, M., & Paulus, M. (2025). Offering and showing gestures in 12- to 15-month-old infants in natural contexts: A corpus-based study. *European Journal of Developmental Psychology*, 22(3), 395–417. [CrossRef]
- Harmati-Pap, V., Vadász, N., Tóth, I., & Kas, B. (2025). Patterns of lexical and syntactic adjustment in early infant-directed speech related to language development in Hungarian. *Clinical Linguistics & Phonetics*, 39(6–8), 736–764.
- Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language*, 44(5), 1140–1162. [PubMed]
- Háden, G. P., Mády, K., Török, M., & Winkler, I. (2020). Newborn infants differently process adult directed and infant directed speech. *International Journal of Psychophysiology*, 147, 107–112. [CrossRef] [PubMed]
- Hilton, C. B., Moser, C. J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C. M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M. B., Martin, A., Cirelli, L. K., Trehub, S. E., Song, J., Kim, M., ... Mehr, S. A. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, 6(11), 1545–1556. [CrossRef] [PubMed]
- Iverson, J. M., Capirci, O., Longobardi, E., & Caselli, M. C. (1999). Gesturing in mother-child interactions. *Cognitive Development*, 14(1), 57–75. [CrossRef]
- Kalashnikova, M., & Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *Journal of Child Language*, 45(5), 1035–1053. [CrossRef] [PubMed]
- Kandemir, S., Özer, D., & Aktan-Erciyes, A. (2024). Multimodal language in child-directed versus adult-directed speech. *Quarterly Journal of Experimental Psychology*, 77(4), 716–728.
- Kang, J. Y., Kim, Y.-S., & Pan, B. A. (2009). Five-year-olds' book talk and story retelling: Contributions of mother—Child joint bookreading. *First Language*, 29(3), 243–265. [CrossRef]
- Kao, C., Sera, M. D., & Zhang, Y. (2022). Emotional speech processing in 3- to 12-month-old infants: Influences of emotion categories and acoustic parameters. *Journal of Speech, Language, and Hearing Research*, 65(2), 487–500. [CrossRef] [PubMed]
- Karmazyn-Raz, H., & Smith, L. B. (2023). Discourse with few words: Coherence statistics, parent-infant actions on objects, and object names. *Language Acquisition*, 30(3–4), 211–229. [PubMed]
- Kempe, V., Ota, M., & Schaeffler, S. (2024). Does child-directed speech facilitate language development in all domains? A study space analysis of the existing evidence. *Developmental Review*, 72, 101121. [CrossRef]
- Kendon, A. (1980). Gesticulation and speech: Two aspects of the. *The Relationship of Verbal and Nonverbal Communication*, 25, 207. [CrossRef]
- Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326–347. [CrossRef]
- Kita, S. (2003). *Pointing: Where language, culture, and cognition meet*. Psychology Press.
- Kitamura, C., & Notley, A. (2009). The shift in infant preferences for vowel duration and pitch contour between 6 and 10 months of age. *Developmental Science*, 12(5), 706–714. [CrossRef] [PubMed]
- Kohári, A., Reichel, U. D., Mády, K., Lakatos, K., & Balázs, A. (2024). A dajkanyelv akusztikájának és a csecsemők temperamentumának vizsgálata. *Alkalmazott Nyelvtudomány*, 24, 42–60.
- Kosie, J. E., & Lew-Williams, C. (2024). Infant-directed communication: Examining the many dimensions of everyday caregiver-infant interactions. *Developmental Science*, 27(5), e13515. [CrossRef] [PubMed]
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684–686. [CrossRef] [PubMed]
- Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, 41(3), 841–849. [CrossRef] [PubMed]
- Loban, W. (1976). *Language development: Kindergarten through grade twelve*. National Council of Teachers of English.
- Mangiafico, S. S. (2016). *Summary and analysis of extension program evaluation in R*. Rutgers Cooperative Extension. Available online: <https://rcompanion.org/documents/RHandbookProgramEvaluation.pdf> (accessed on 16 February 2026).
- ManyBabies Consortium. (2020). Quantifying sources of variability in infancy research using the infant-directed-speech preference. *Advances in Methods and Practices in Psychological Science*, 3(1), 24–52. [CrossRef]
- McArthur, D., Adamson, L., & Deckner, D. F. (2005). As stories become familiar: Mother-child conversations during shared reading. *Merrill-Palmer Quarterly*, 51(4), 389–411. [CrossRef]

- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.
- Mengru, H., & Kager, R. (2024). Relating the prosody of infant-directed speech to children's vocabulary size. *Journal of Child Language*, 51(1), 217–233.
- Miyazawa, K., Shinya, T., Martin, A., Kikuchi, H., & Mazuka, R. (2017). Vowels in infant-directed speech: More breathy and more variable, but not clearer. *Cognition*, 166, 84–93. [CrossRef] [PubMed]
- Murphy, C. M. (1978). Pointing in the context of a shared activity. *Child Development*, 49(2), 371–380. [CrossRef]
- 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. [CrossRef] [PubMed]
- Nicoladis, E., Nagpal, J., Marentette, P., & Hauer, B. (2018). Gesture frequency is linked to story-telling style: Evidence from bilinguals. *Language and Cognition*, 10(4), 641–664. [CrossRef]
- Parlato-Oliveira, E., Chetouani, M., Cadic, J.-M., Viaux, S., Ghattassi, Z., Xavier, J., Ouss, L., Feldman, R., Muratori, F., Cohen, D., & Saint-Georges, C. (2020). The emotional component of Infant Directed-Speech: A cross-cultural study using machine learning. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, 68(2), 106–113. [CrossRef]
- Pellegrino, E., Dellwo, V., Pardo, J. S., & Möbius, B. (2024). *Forms, factors and functions of phonetic convergence* (Vol. 165). Elsevier.
- Pickering, M. J., & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation*, 4(2), 203–228. [CrossRef]
- Pika, S., Nicoladis, E., & Marentette, P. F. (2006). A cross-cultural study on the use of gestures: Evidence for cross-linguistic transfer? *Bilingualism: Language and Cognition*, 9(3), 319–327. [CrossRef]
- Polka, L., Masapollo, M., & Ménard, L. (2022). Setting the stage for speech production: Infants prefer listening to speech sounds with infant vocal resonances. *Journal of Speech, Language, and Hearing Research*, 65(1), 109–120. [CrossRef] [PubMed]
- Prieto, P., Esteve-Gibert, N., & Shattuck-Hufnagel, S. (2025). Towards a novel conceptualization of prosody that accounts for spoken and visual signals: The modality-neutral prosodic framework hypothesis. *Gesture*, 23(1–2), 119–159.
- 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. [CrossRef] [PubMed]
- R Core Team. (2025). *R: A language and environment for statistical computing* [Computer software manual]. Available online: <https://www.R-project.org/> (accessed on 31 October 2025).
- Reichel, U. D. (2016). *CoPaSul manual—contour-based parametric and superpositional intonation stylization* [Computer software manual]. Available online: <https://github.com/reichelu/copasul> (accessed on 25 August 2022).
- Rohlfing, K. J., Grimminger, A., & Nachtigäller, K. (2015). Gesturing in joint book reading. In *Learning from picturebooks* (pp. 99–116). Routledge.
- Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, 15(3), 426–435. [CrossRef] [PubMed]
- Rosslund, A., Mayor, J., Mundry, R., Singh, A. P., Cristia, A., & Kartushina, N. (2024). A longitudinal investigation of the acoustic properties of infant-directed speech from 6 to 18 months. *Royal Society Open Science*, 11(11), 240572. [CrossRef] [PubMed]
- Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., Laznik, M.-C., & Cohen, D. (2013). Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS ONE*, 8(10), e78103. [CrossRef] [PubMed]
- Salo, V. C., Reeb-Sutherland, B., Frenkel, T. I., Bowman, L. C., & Rowe, M. L. (2019). Does intention matter? Relations between parent pointing, infant pointing, and developing language ability. *Journal of Cognition and Development*, 20(5), 635–655. [CrossRef] [PubMed]
- Schmidt, C. L. (1996). Scrutinizing reference: How gesture and speech are coordinated in mother-child interaction. *Journal of Child Language*, 23(2), 279–305. [CrossRef] [PubMed]
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–671. [CrossRef] [PubMed]
- Smith, N. A., & Strader, H. L. (2014). Infant-directed visual prosody: Mothers' head movements and speech acoustics. *Interaction Studies*, 15(1), 38–54. [CrossRef] [PubMed]
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27(4), 501–532. [CrossRef]
- Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, 44, 1–18. [CrossRef]
- Stavans, A. (1996). Development of parental narrative input. *Journal of Narrative and Life History*, 6(3), 253–280. [CrossRef]
- Stavans, A., & Goldzweig, G. (2008). Parent-child-adult storytelling: Commonalities, differences and interrelations. *Narrative Inquiry*, 18(2), 230–257. [CrossRef]
- Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, 17(4), 432–444. [PubMed]
- 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. [CrossRef]

- Tóth, B., Háden, G., Tóth, I., Lakatos, K., Kohári, A., Mády, K., Kas, B., Tóth, D., Szalontai, Á., Reichel, U. D., & Winkler, I. (2025). Predicting infant vocabulary from neural connectivity and maternal speech: A machine learning approach. *Brain Research*, 1866, 149892. [CrossRef] [PubMed]
- van Schaik, J. E., Meyer, M., van Ham, C. R., & Hunnius, S. (2020). Motion tracking of parents' infant-versus adult-directed actions reveals general and action-specific modulations. *Developmental Science*, 23(1), e12869. [PubMed]
- Vigliocco, G., Motamedi, Y., Murgiano, M., Wonnacott, E., Marshall, C., Maillo, I. M., & Perniss, P. (2019). Onomatopoeia, gestures, actions and words: How do caregivers use multimodal cues in their communication to children? In *Proceedings of the annual meeting of the cognitive science society* (Vol. 41). Cognitive Science Society.
- Wei, T., & Simko, V. (2024). *R package 'corrplot': Visualization of a correlation matrix* (Version 0.95) [Computer software manual]. Available online: <https://github.com/taiyun/corrplot> (accessed on 16 February 2026).
- Wigglesworth, G., & Stavans, A. (2001). A crosscultural investigation of Australian and Israeli parents' narrative interactions with their children. In *Children's language* (pp. 73–91). Psychology Press.
- Xu, N., & Burnham, D. (2010, May 11–14). *Tone hyperarticulation and intonation in Cantonese infant directed speech*. Speech Prosody 2010 (p. 094), Chicago, IL, USA. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.