


Acoustic Properties of Infant Directed Speech are Similar in Video Chat Versus In-Person Interactions


Acoustic Properties of Infant Directed Speech are Similar in Video Chat Versus In-Person Interactions
 Madeline R. Hale and Melanie J. Spence
 School of Behavioral and Brain Sciences, The University of Texas at Dallas

Introduction

- Caregivers speak to infants using infant-directed speech (IDS) and infants respond to IDS with gaze and affective states.
- IDS differs from adult-directed speech (ADS) by acoustic properties of high and variable fundamental frequency (mean F_0 and F_0 range, respectively), short utterances, long pauses, and vocal inflections, which collectively contribute to the parents' or mothers' perception of "sing-song".

OPEN

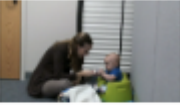
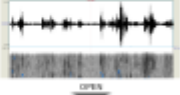
Methods

- 28 mothers and their 2-month-old infants participated in free play and video chat during a lab visit.
- 17% of participants reported their ethnicity as Caucasian, 18% as Hispanic, 7% as Asian, and 25% as African American.

Pitch Analysis

- Acoustic samples from 28 mothers were collected to derive segmental statistics:

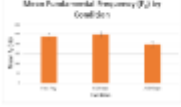
1. Mean F_0 range (mean of mothers and their 2-month-old infants)


OPEN

Results

Mean Fundamental Frequency (F_0) by Condition



Mean Frequency (F_0) Range by Condition



Descriptive Statistics

	Mean- F_0 (Hz)	F_0 Range (Hz)
Free-Play	239.57 (8.6% SD)	345.70 (15.98 SD)

OPEN

Conclusion

- The mean F_0 and F_0 range values for free play and video chat differed from those collected for ADS, as is typical in the literature.
- No significant acoustic differences were found between mothers' IDS for in-person versus video interactions.
- This suggests that adults' reports of a "sing-song" style of speech, some of the salient prosodic properties of IDS that caregivers use to engage infants, might be more consistent than we first believe.
- This finding extends our knowledge that video chat may affect infant attention and... (text partially obscured)

OPEN

Research Question

Does the change in context between in-person versus video chat interactions affect the acoustic properties of mothers' IDS?

References & Acknowledgments

References

1. Fernald, A., Tanenhaus, T., Diaz, J., Papenow, M., DeKeyser-Blake, S., & Pádua, I. (2002). A cross-linguistic study of prosodic modifications in mothers' and fathers' speech to preschool children. *Journal of Child Language*, 29(3), 477-503.
2. Fernald, A. (2002). Attentional/functional in mothers' speech to infants. In R. Papenow, D. Zayas, & M. Papenow (Eds.), *Infants' attention and vocal interaction: Molecular/early communication: Communication and developmental*

[CHAT SESSION](#)
[NARRATION](#)
[AUTHOR INFORMATION](#)
[ABSTRACT](#)
[REFERENCES](#)
[CONTACT AUTHOR](#)
[PRINT](#)

Madeline R. Hale and Melanie J. Spence

School of Behavioral and Brain Sciences, The University of Texas at Dallas



PRESENTED AT:





Society for Research in Child Development

2021 *Virtual* BIENNIAL MEETING

APRIL 7-9, 2021

INTRODUCTION

- Caregivers speak to infants using infant-directed speech (IDS),¹ and infants respond to IDS with positive affect and attention^{4,9}.
- IDS differs from adult-directed speech (ADS) by acoustic properties of high and variable fundamental frequency (mean- F_0 and F_0 range, respectively), short utterances, long pauses, and vowel alterations, which collectively contribute to the prosody or melodic properties of speech^{1,2}.
- The prosody of IDS communicates caregivers' emotional intent^{3,6}, affects infants' emotional regulation^{2,9}, and recruits infants' attention^{4,5}.
- Video chat (Skype, FaceTime) has become a widely used technology for infants to develop relationships and communicate with long-distance family members¹⁰ and this widespread use has only increased due to the COVID-19 pandemic¹¹.
- Studies have explored infant attention and learning over video chat,^{7,12} but studies have not investigated if the acoustic properties of IDS are similar in face-to-face versus video chat interactions.

RESEARCH QUESTION

Does the change in context between in-person versus video chat interactions affect the acoustic properties of mothers' IDS?

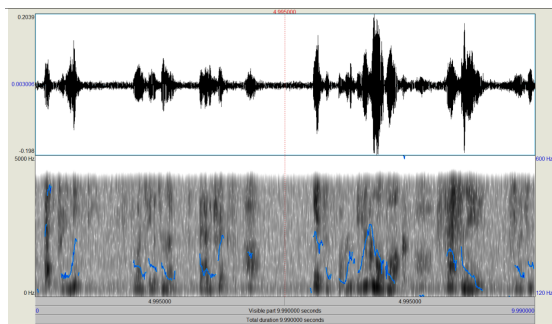
METHODS

Participants

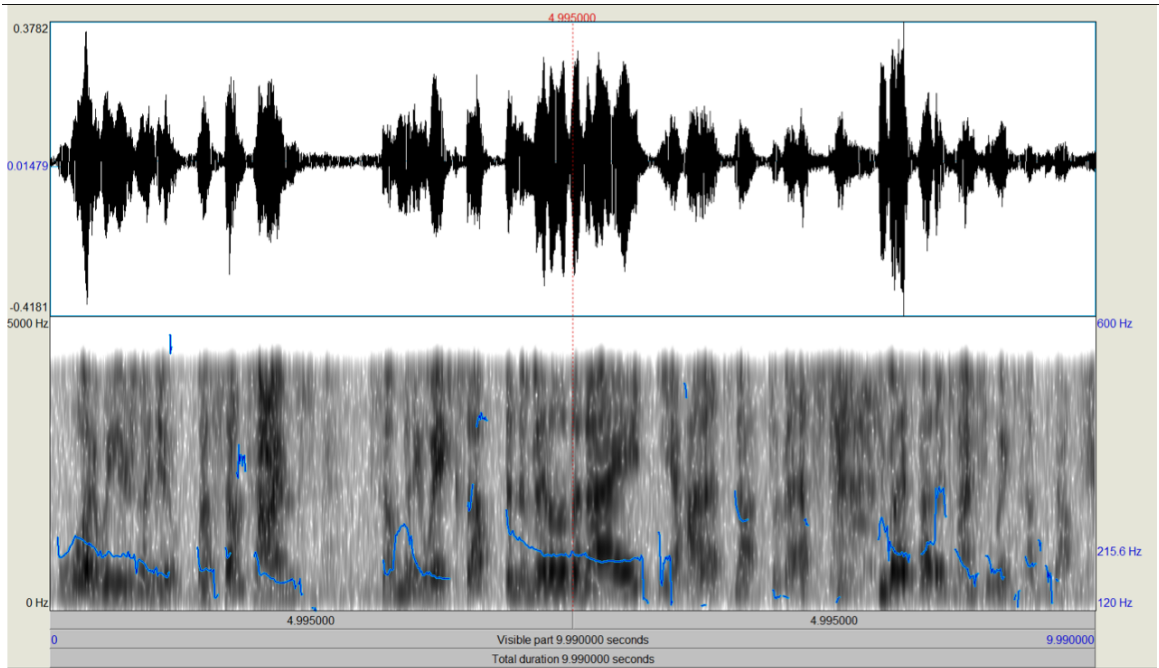
- 29 mothers and their 5-month-old infants participated in free-play and video chat during a lab visit.
- 79% of participants reported their ethnicity as Caucasian, 10% as Hispanic, 7% as Asian, and 4% as African American.

Pitch Analysis

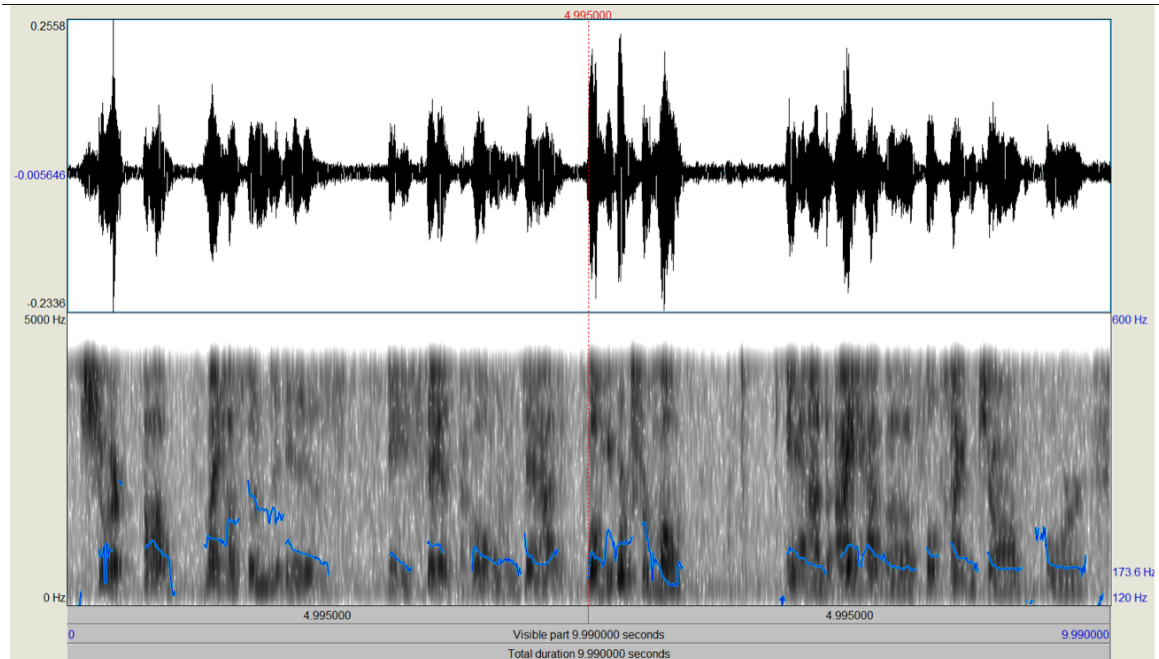
- Audio samples from 29 mothers were collected in three experimental contexts:
 1. 3-min **free-play** interaction of mothers with their 5-month-old infant



2. A Skype session between the same dyad discussing their weekly schedule (**IDS**)



3. A Skype conversation between the mother and an adult research assistant explaining her weekly schedule (ADS)



- Mother-infant free-play and Skype interactions were converted into WAV files using VLC Media Player.
- The first 10-s (after a 10-s warm-up period) of “uninterrupted vocalizations” from the mother in each of the three conditions were chosen for analysis according to the criteria established in Broesch and Bryant (2015).
- The mean- F_0 and F_0 range of the mother’s speech in three situations was processed using Praat Version 6.0.48.
 - Mean- F_0 was computed from the average of all F_0 stimuli Praat recognized in the sample.
 - F_0 range was computed by taking the difference between the highest and lowest F_0 s.

Analyses

- Dependent Measures: Mean- F_0 and F_0 range derived from 10s of maternal speech
- Two Repeated Measures Anovas compared each dependent variable in the three contexts.

Anova 1:
Mean- F_0

Free-Play
Mean- F_0

x

IDS Skype
Mean- F_0

x

ADS Skype
Mean- F_0

Anova 2:
 F_0 Range

Free-Play
 F_0 Range

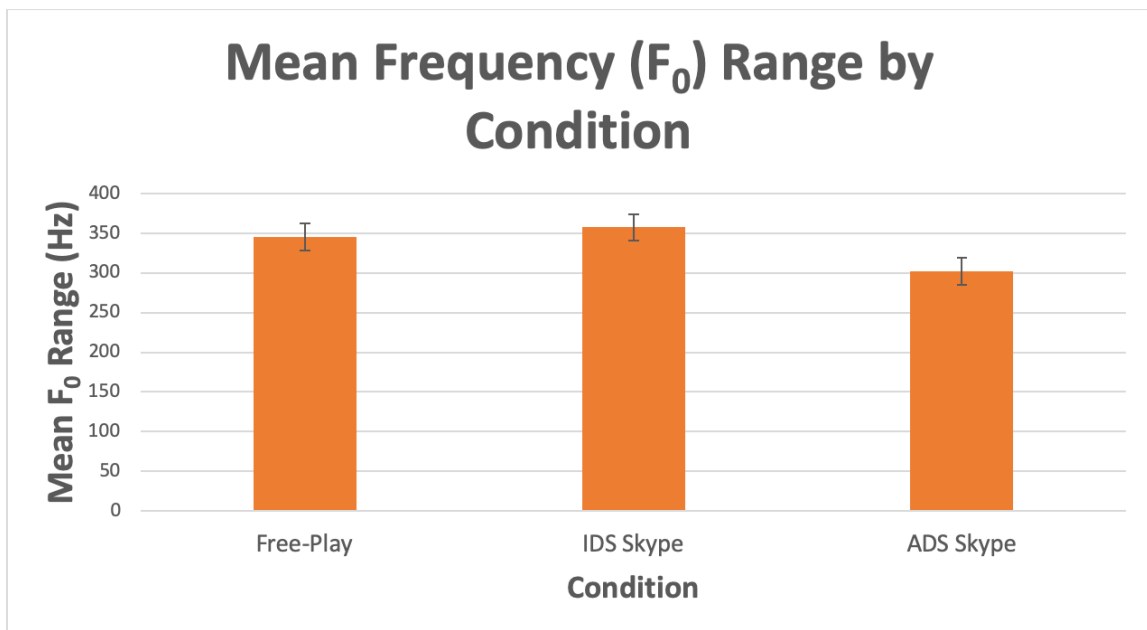
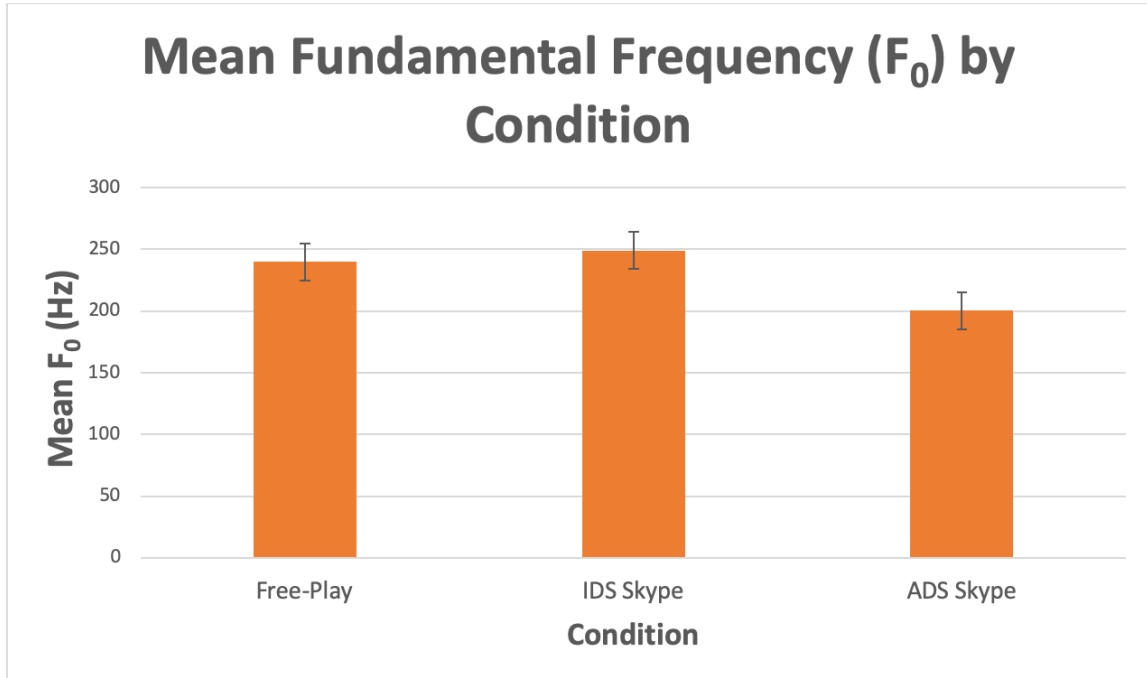
x

IDS Skype
 F_0 Range

x

ADS Skype
 F_0 Range

RESULTS



Descriptive Statistics

	Mean-F₀ (Hz)	F₀ Range (Hz)
Free- Play	239.87 (5.58 SE)	345.76 (13.89 SE)
IDS Skype	248.98 (5.57 SE)	357.48 (12.97 SE)
ADS Skype	200.41 (2.256 SE)	301.79 (12.43 SE)

Repeated Measures Anova 1

Main Effect Mean-F₀

F(1.38,38.86) = 25.70, p<.001

Greenhouse-Geisser Correction

Post Hoc Pairwise Comparison:

Mean-F₀	
Free-Play vs ADS Skype	$p < .001$
IDS Skype vs ADS Skype	$p < .001$
Free-Play vs IDS Skype	$p = 1$

Repeated Measures Anova 2

Main Effect F₀ Range

F(2,56)=7.14, p=.002

Post Hoc Pairwise Comparison:

F₀ Range	
Free-Play vs ADS Skype	$p < .001$
IDS Skype vs ADS Skype	$p < .001$
Free-Play vs IDS Skype	$p = 1$

CONCLUSION

- The mean- F_0 and F_0 range values for free-play and Skype IDS differed from those collected for ADS, as is typical in the literature¹.
- No significant acoustic differences were found between mothers' IDS for in-person versus Skype interactions.
- This suggests that adults' speech to infants over video chat conveys some of the salient prosodic properties of IDS that communicate caregiver affect, regulate infants' emotion, and direct infant attention.
- This finding extends our knowledge that video chat can affect infant attention and learning to providing specific data on the actual input infants receive^{7,12}.
- Further work is needed to explore additional characteristics of IDS prosody in video chat to fully elucidate its possible impacts on infants' socio-emotional development.
- Limitations of this study included a relatively small sample size that does not represent the racial make-up of the United States.

REFERENCES & ACKNOWLEDGMENTS

References

1. 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.
2. Fernald, A. (1992). *Meaningful melodies in mothers' speech to infants*. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Studies in emotion and social interaction. Nonverbal vocal communication: Comparative and developmental approaches* (p. 262–282). Cambridge University Press; Editions de la Maison des Sciences de l'Homme.
3. Bryant, G.A., & Barrett, H.C. (2007). Recognizing Intentions in Infant-Directed Speech: Evidence for Universals. *Psychological Science*, 18(8), 746-751.
4. Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 43(2), 230–246.
5. Kaplan, P. S., Bachorowski, J.-A., Smoski, M. J., Hudenko, W. J., & Zarlengo-Strouse, P. (1995). Habituation, sensitization, and infants' responses to motherese speech. *Developmental Psychobiology*, 28, 45–57.
6. Kitamura, C., & Lam, C. (2009). Age-Specific Preferences for Infant-Directed Affective Intent. *Infancy*, 14(1), 77-100.
7. Roseberry, S., Hirsh-Pasek, K. and Golinkoff, R.M. (2014). Skype Me! Socially Contingent Interactions Help Toddlers Learn Language. *Child Development*, 85, 956-970.
8. 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.
9. Fernald, A. (1993). Approval and Disapproval: Infant Responsiveness to Vocal Affect in Familiar and Unfamiliar Languages. *Child Development*, 64(3), 657-674.
10. McClure E., & Barr R. (2017). Building Family Relationships from a Distance: Supporting Connections with Babies and Toddlers Using Video and Video Chat. In Barr R. & Linebarger D. (Eds.), *Media Exposure During Infancy and Early Childhood* (pp. 227-248). Springer, Cham.
11. Wiederhold, B.K. (2020). Children's Screen Time During the COVID-19 Pandemic: Boundaries and Etiquette. *Cyberpsychology, Behavior, and Social Networking*, 23(6), 359-360.
12. McClure, E.R., Chentsova-Dutton, Y.E., Holochwost, S.J., Parrott, W.G., Barr, R. (2018). Look At That! Video Chat and Joint Visual Attention Development Among Babies and Toddlers. *Child Development*, 89, 27-36.

Acknowledgments

Special thanks to all the families who participated in our study, and to the research assistants at the Infant Learning Project.

Thank you to Dr. Kate Shepard for the collection of video and Skype data.

AUTHOR INFORMATION

Madeline Hale is an undergraduate senior at the University of Texas at Dallas double majoring in Speech-Language Pathology & Audiology and Neuroscience. Her work with Dr. Melanie Spence in the Infant Learning Project has focused on the content and acoustic properties of infant-directed speech. This exploratory project formed the basis of her honors thesis that examines this same research question with additional participants and acoustic measures. In the fall, Madeline will begin her graduate studies in speech-language pathology.

Melanie Spence is Professor of Psychology at The University of Texas at Dallas (UTD) and serves as the Associate Dean for Undergraduate Education for the School of Behavioral and Brain Sciences at UTD. She received her Ph.D. from the University of North Carolina at Greensboro. She has been recognized as a Fellow of the Association for Psychological Science.

ABSTRACT

Infant-directed speech (IDS) differs from adult-directed speech (ADS) in acoustic properties of high and variable frequency (mean- F_0 and F_0 range, respectively), short utterances, and vowel alterations, which collectively contribute to the prosody of speech (Fernald et al., 1989; Kitamura et al., 2001). Salient acoustic characteristics of IDS have been linked to processes important for language acquisition, such as speech segmentation (Golinkoff et al., 2015). Additionally, IDS prosody plays a role in communication of caregivers' positive emotions and regulation of infants' emotional states during the first year (Fernald, 1992; Bryant & Barrett, 2007). For example, infants respond with greater positive affect to IDS than ADS (Werker and McLeod, 1989) and infants' preference for the emotional content of IDS differs across development (Kitamura et al., 2009), demonstrating caregiver-infant emotional communication in IDS.

In this age of technology, heightened by the COVID-19 pandemic, family members often communicate to infants over video chat. Research has not investigated whether IDS prosody varies significantly in face-to-face versus virtual interactions. The current study examines the differences in acoustic properties of IDS over Skype versus in-person play. Specifically, we analyzed the mean- F_0 and F_0 range (minimum- F_0 subtracted from maximum- F_0) of 29 mothers in three conditions: (1) a 3-min free-play interaction with their 5-month-old infant, (2) a Skype session with their infant, and (3) a Skype conversation with an adult researcher. In Skype situations mothers were instructed to discuss their weekly schedule.

Videos of mother-infant free-play and Skype interactions were converted into WAV files using VLC Media Player. A 10-s warm-up period was given at the beginning of each recording to allow mothers to adjust to the lab environment. The next 10-s of uninterrupted vocalizations from the mother were analyzed in Praat version 6.0.48 according to the criteria established in Broesch and Bryant (2013).

The mean- F_0 and F_0 range that typically distinguish IDS from ADS were consistent with those previously reported (Fernald, et al., 1989). Separate within-subject Anovas comparing mean- F_0 and F_0 range across the three situations resulted in a significant main effect for mean- F_0 ($F(1.38,38.86) = 25.70, p < .001$, Greenhouse-Geisser correction) and for F_0 range ($F(2,56)=7.14, p=.002$). Post hoc pairwise comparisons revealed significant differences for mean- F_0 between free-play ($M=239.87, SD=30.09$) and ADS Skype ($M=200.40, SD=13.76$) ($p < .001$) and between IDS Skype ($M=248.98, SD=29.99$) and ADS Skype ($p < .001$), but not between free-play and IDS Skype ($p=1.0$). The same group differences were found for F_0 range; free-play ($M=345.76, SD=74.78$) and IDS Skype ($M=357.48, SD=69.84$) did not differ ($p=1.0$), while both differed from ADS Skype ($M=301.79, SD=66.92$) (free-play vs. ADS $p=.027$, IDS Skype vs. ADS $p=.001$). These results suggest that adults' speech to infants over video chat conveys some of the salient prosodic properties of IDS that communicate and regulate emotion. They also extend knowledge of infant word learning from video chat interactions to the prosodic properties of the speech itself (Roseberry et al., 2014). Future directions should investigate whether other acoustic properties of IDS, such as slowed speaking rate and vowel hyper-articulation, are present in video chat interactions.

REFERENCES

1. 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.
2. Fernald, A. (1992). *Meaningful melodies in mothers' speech to infants*. In H. Papoušek, U. Jürgens, & M. Papoušek (Eds.), *Studies in emotion and social interaction. Nonverbal vocal communication: Comparative and developmental approaches* (p. 262–282). Cambridge University Press; Editions de la Maison des Sciences de l'Homme.
3. Bryant, G.A., & Barrett, H.C. (2007). Recognizing Intentions in Infant-Directed Speech: Evidence for Universals. *Psychological Science*, 18(8), 746-751.
4. Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 43(2), 230–246.
5. Kaplan, P. S., Bachorowski, J.-A., Smoski, M. J., Hudenko, W. J., & Zarlengo-Strouse, P. (1995). Habituation, sensitization, and infants' responses to motherese speech. *Developmental Psychobiology*, 28, 45–57.
6. Kitamura, C., & Lam, C. (2009). Age-Specific Preferences for Infant-Directed Affective Intent. *Infancy*, 14(1), 77-100.
7. Roseberry, S., Hirsh-Pasek, K. and Golinkoff, R.M. (2014). Skype Me! Socially Contingent Interactions Help Toddlers Learn Language. *Child Development*, 85, 956-970.
8. 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.
9. Fernald, A. (1993). Approval and Disapproval: Infant Responsiveness to Vocal Affect in Familiar and Unfamiliar Languages. *Child Development*, 64(3), 657-674.
10. McClure E., & Barr R. (2017). Building Family Relationships from a Distance: Supporting Connections with Babies and Toddlers Using Video and Video Chat. In Barr R. & Linebarger D. (Eds.), *Media Exposure During Infancy and Early Childhood* (pp. 227-248). Springer, Cham.
11. Wiederhold, B.K. (2020). Children's Screen Time During the COVID-19 Pandemic: Boundaries and Etiquette. *Cyberpsychology, Behavior, and Social Networking*, 23(6), 359-360.
12. McClure, E.R., Chentsova-Dutton, Y.E., Holochwost, S.J., Parrott, W.G., Barr, R. (2018). Look At That! Video Chat and Joint Visual Attention Development Among Babies and Toddlers. *Child Development*, 89, 27-36.