Toddlers Use Vocal Cues to Infer Male Dominance Over Females in Right-of-Way Conflict

Flavia Filesi[1] (flavia.filesi@psykologi.uio.no), Bjørn Dahl Kristensen[1] (b.d.kristensen@psykologi.uio.no), Erik Kjos Fonn[1] (e.k.fonn@psykologi.uio.no), Joakim Haugane Zahl[1] (j.h.zahl@psykologi.uio.no), Mathilde Massenet[3] (massenet.mathilde@gmail.com), Gregory A. Bryant[3] (gabryant@ucla.edu), and Lotte Thomsen[1,2](lotte.thomsen@psykologi.uio.no)

[1]Department of Psychology, University of Oslo,
Forskningsveien 3A, 0373, Oslo, Norway
[2]Department of Political Science, Aarhus University
Bartholins Allé 7, 8000 Aarhus C, Denmark
[3]Department of Communication, Center for Behavior, Evolution, and Culture, University of California,
Los Angeles, California 90095, USA

Abstract

Men hold more power compared to women, a phenomenon that is globally evident and widely documented. Preverbal infants mentally represent social dominance, and by the age of 18 months can distinguish male and female voices and associate them with faces of the corresponding gender. Here, we investigated whether 18- to 24- month-old toddlers (N=48) expected male-voiced agents to prevail over female-voiced ones in a right-of-way conflict. Using a violation-of-expectation paradigm, we found that toddlers looked longest when female-voiced characters prevailed in dominance conflicts, suggesting they expected that male-voiced agents would prevail instead. Acoustic features of male voices that perceptually suggest greater formidability/physical size (i.e., lower fundamental frequency) may account for this effect. Alternatively, vocal indicators of speaker sex might trigger very early conceptualizations about gender and dominance.

Keywords: Dominance; Gender; Voice; Formant Frequency; Fundamental Frequencies

Introduction

Men tend to occupy positions of power and advantage over women, a pattern observable across multiple domains of social life and remarkably consistent across cultures and historical periods. Female economic participation remains significantly lower than that of men, and women are underrepresented in leadership positions and in the political sphere (see Global Gender Gap Report 2024 by the World Economic Forum). It is not surprising, therefore, that we tend to make implicit associations assuming that men have higher status than women (Salles et al., 2019; Hansen et al., 2019).

Gender differences in dominance motives are also well-documented in social and political psychology with males having higher average social dominance orientation levels than females (Lee, Pratto, & Johnson, 2011; Kunst et al., 2017). Gender differences related to social dominance also arise in paraverbal communication where men tend to interrupt more frequently and use direct commands and assertive tones, while women tend to favor indirect requests and more cautious language (Anderson & Leaper, 1998). Furthermore, men adopt open and dominant postures, speak at louder volumes, and maintain direct eye contact while speaking, whereas women tend to display more closed postures and greater eye contact while listening (Lakoff, 1973; Hall et al., 2005; Dovidio et al., 1988).

Critically, the physical characteristics that are more typical of males—such as being taller, having muscular bodies, deep voices, and more dominant facial features (larger jaws, wider chins, and thicker brows)—robustly influence perceived dominance. This is true for males as well as females, although these characteristics are more typical of males both in humans and in most other primates (Lukaszewski et al., 2016; Murray &, Schmitz, 2011; Muller & Mazur, 1997; Windhager, Schaefer, & Fink, 2011, 2011; Toscano, Schubert, & Sell, 2014; Plavcan, 2001).

Young children seem to represent these patterns very early. By ten months of age, they are sensitive to the relationship between physical size and dominance, expecting larger agents to prevail over smaller ones in zero-sum conflict (Thomsen et al., 2011). The association between masculine facial features, and perceived strength (i.e. formidability), and dominance has been demonstrated in preschoolers (Cogsdill et al., 2014; Keating & Bai, 1986; Terrizzi et al., 2019). At this age, conversational sex differences are also observed, with boys interrupting girls at a 2:1 ratio (Esposito, 1979). The association between males and power has been observed in children as young as four years old during mixed-gender interactions. This association was found both when power was represented through body postures and when it involved control over decision-making and resources (Charafeddine et al., 2020). Whether infants and young toddlers also associate being male with greater dominance, however, remains unknown.

The ability to recognize gender-specific vocal traits emerges early in life. By six months of age, infants begin to develop the ability to match female faces with female voices and male faces with male voices (Miller, Younger, & Morse, 1982; Walker-Andrews et al., 1991; Patterson & Werker, 2002). Evidence suggests that by one year of age, this ability

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

is generally well developed. However, findings vary depending on the experimental paradigm used: while some studies report successful matching for both male and female stimuli, others show that infants succeed only when female voices and faces are involved (Poulin-Dubois et al., 1994; Hillairet de Boisferon et al., 2015). By 18 months, however, this ability appears firmly established across genders (Poulin-Dubois, Serbin, & Derbyshire, 1998). These studies do not identify which vocal cues children rely on to distinguish between female and male voices, but men's and women's voices exhibit typical average differences in certain acoustic parameters, specifically fundamental frequency (f_o) and formant frequencies (Titze, 1994).

Fundamental frequency (f_o) is the acoustic correlate of perceived voice pitch and is determined by the rate of vocal fold vibrations. Longer and thicker vocal folds vibrate more slowly, resulting in a lower f_o and pitch (Titze, 1994; Fitch, 1997). During puberty, testosterone levels influence the growth of the vocal folds and the larynx, as androgen receptors are abundant in vocal fold cells (Titze, 1994; Harries et al., 1997, 1998; Zamponi et al., 2021). This hormonal influence creates pronounced sexual dimorphism in f_o: men's vocal folds are approximately 60% longer than women's, leading to an average pitch about an octave lower (Titze, 1994; Puts, Apicella, & Cárdenas, 2012). Formant frequencies are the resonant frequencies of the vocal tract (Fant, 1960). Longer vocal tracts produce lower and more closely spaced formants. Since men generally have longer vocal tracts than women, their formant frequencies are also lower on average (Hillenbrand et al., 1995).

Research has shown that low f_o and formant frequencies influence perceptions of speakers as being larger, stronger, and better leaders. This applies to both male and female voices. Formants are a more reliable predictor of body size than f_o (Fitch, 2000), which shows only a very weak correlation with body size in adult humans and other mammals (Künzel, 1989; McComb, 1991; Reby & McComb, 2003; Rendall et al., 2005; Taylor & Reby, 2010). Regardless of the actual relationship with physical size, low f_o and formants consistently create the impression of greater physical body size and dominance. This bias is prevalent across the animal kingdom as well as within our own species (Pisanski et al., 2014; Rendall, Vokey, & Nemeth, 2007; Pietraszewski et al., 2017; Vannoni & McElligott, 2008; Taylor, Reby & McComb, 2010). In addition to appearing larger, men with lower-pitched voices are also perceived as stronger, older, more dominant, and more masculine (Aung & Puts, 2020). Interestingly, humans also modulate their voices by lowering their pitch in social contexts where they want to assert dominance and by raising it (submissively) when addressing individuals who outrank them (Puts, Gaulin, & Verdolini, 2006; Leongómez et al., 2017; Zhang et al., 2021).

Research has also shown that vocal pitch influences perceptions of leadership and competence. Individuals tend to lower their pitch when they want to appear more competent or authoritative, which effectively leads listeners to rate them this way (Sorokowski et al., 2019). We tend to prefer leaders—both male and female—with lower-pitched voices, because we perceive them as more competent and trustworthy (Klofstad, Anderson, & Peters, 2012; Klofstad, Anderson, & Nowicki, 2015). Strikingly, lower-pitched voices (i.e., more masculine voices) are preferred even for traditionally feminine leadership roles (Anderson, & Klofstad, 2012).

In sum, we tend to expect deeper voices to belong to individuals who are physically larger, stronger, and more powerful—all aspects which cue the core concept of social dominance (Thomsen et al., 2011; Thomsen, 2020; Mascaro & Csibra, 2012; Gazes, Hampton, & Lourenco, 2017). Hence, infants and toddlers may already be biased towards perceiving male voices as more dominant than female voices, which tend to be higher pitched. If so, this would present a mechanism, grounded in the biology of male and female sex-typed voices, that license the general inference that females will submit to males in dominance conflict, a theme recurrent in gendered stereotypes and roles.

Here, we investigated whether 18- to 24-month-old children expected an agent with a male voice to prevail in a dominance conflict over an agent with a female voice. To do so, we modified the well-validated right-of-way dominance paradigm in which two novel agents block each other's path before one of them submits by prostrating and yielding the right-of-way. Using this paradigm, Thomsen et al. (2011) showed that preverbal infants expected larger, more formidable individuals to prevail in zero-sum conflict over smaller ones; Pun, Birch, and Baron (2016, 2022) showed that children expected a member of a larger coalition to prevail over a member of smaller one, and Thomas et al. (2018) showed that toddlers prefer those (same-sized) agents who dominated in a right-of-way conflict. Here, we used male and female voices to cue the relative formidability of two same-sized agents.

Methods

Experimental design

Following the procedures of Thomsen et al, (2011), toddlers were shown animations of agents interacting, following the procedures established by Thomsen et al. (2011). During familiarization (see male voice familiarization https://bit.ly/3CpZRSZ; see female voice familiarization http://bit.ly/4hIj0y6. Duration: 19 seconds), each agent in turn crosses a platform on their own, in opposite directions. This serves to show that each agent has the goal of moving to the opposite side of the platform from where it started. Next, in an intertrial, both agents simultaneously move from their habitual beginning positions so that they meet in the middle, blocking each other's way, before withdrawing (see intertrial https://bit.ly/4gu7ndj. Duration: 27 seconds). This serves to highlight the zero-sum conflict between their goals. Two test events follow. In the *Expected outcome* test trial, the female-voiced agent prostates before the other and then scoots sideways out of the way, upon which the male-voice agent continues to complete its goal of crossing the stage (https://bit.ly/4gqo0qb. Duration: 38 seconds). In the *Unexpected outcome* test trial, these roles reverse to that female-voiced agent prevails (https://bit.ly/4hD6qQI. Duration: 38 seconds). Thus, the Expected and Unexpected outcomes only differ in whether the prevailing agent has a male or female voice.

Stimuli were animated in Keynote. We also animated the mouth of each agent: it moves in synchrony with the vocalizations to reinforce the perception that the sound is being produced by the agents. The color (orange/purple), placement (left/right), and order of appearance of the male- and female voiced character, together with the presentation order of expected and unexpected test trials, were counter-balanced across participants.

The sounds in the animations were recorded using a Svive Hydra Pro microphone. The natural vocalizations were produced by a 26-year-old woman (female voice) and a 23-year-old man (male voice). Speakers were instructed to produce brief, nonverbal vocal sounds to accompany specific movements in the animations. The recordings were edited using Audacity to remove background noise and improve overall clarity, but no manipulations were made to pitch, duration, or other acoustic features. The vocalizations were fully nonverbal and contained no lexical or word-level content. Table 1 reports the pitch and formant characteristics of each vocalization. On average, the female voices had a higher fundamental frequency (mean F0 = 205 Hz) than the male voices (mean F0 = 128 Hz), as well as higher F1 and F2 formants (F1 = 545 Hz and F2 = 1002 Hz for the female voices, compared to F1 = 402 Hz and F2 = 923 Hz for the male voices), reflecting typical sex-based acoustic differences.

Table 1: Fundamental and formant values (f_o, F1, F2) for each vocalization by speaker gender.

Vocal Sounds	f_o Mean (Hz)	f_o SD (Hz)	F1 (Hz)	F2 (Hz)
F approaches	266	60	548	1148
F backs away	190	12	696	1284
F bows	270	17	600	1065
F jump	267	67	574	1051
M approaches	144	55	460	908
M backs away	89	10	638	1085
M bows	112	15	518	973
M jump	167	46	428	938

Participants

We recruited parents of toddlers who had expressed interest in participating in a study on evolutionary, social, and developmental psychology through an online form. While the majority of participants were from the metropolitan area of Oslo, as the study was conducted digitally, some participants were located in other regions of Norway. The final sample size included 48 children aged 18 to 24 months, with three participants per counter-balance condition (27 girls, 21 boys). An additional 25 participants were excluded for the following reasons: 14 because the minimum looking time was not reached (fussed out or distracted); 2 because parents kept their eyes open during the test phase; 5 due to an error by the experimenter in the presentation of the stimuli; 4 were identified as statistical outliers and replaced.

Procedure

The study was conducted as a video call via Zoom. At the start, the experimenter reviewed the informed consent, and parents verbally consented to their children's participation. The child sat on the parent's lap facing the screen, with the parent's camera feed turned off. This meant that the toddler could not see the parent’s face without turning their head, and neither the child nor the parent could see themselves or the experimenter on the screen. The experimenter then shared her screen to show stimuli, muting her headphones to remain blind to which agent had a male or female voice.

The procedure started with a calibration video, designed to capture the child's attention and establish baseline eye movement parameters based on where the stimulus appeared on the screen. Afterwards, familiarization videos were shown twice, an inter-trial interval occurred, and two alternating test events (Expected and Unexpected outcomes), each shown twice. Before the start of the test phase, a notification appeared on the screen asking the parent to close their eyes. This served to exclude the possibility that children may have inferred anything about the test videos by looking at their parents’ gaze or facial expression. They were also asked to be as neutral as possible during the presentation of the animations and to not encourage their children to look at the screen.

Coding and inclusion criteria

Participants’ looking times for each trial were recorded live by a trained coder using jHab (Casstevens, 2007). For familiarization and intertrial videos, recording began when the agents simultaneously performed their first movement (a jump) and produced their first vocalization. We showed all participants all familiarization trials. In the test trials, coding instead started when one agent initiated the movement and vocalization associated with bowing forward, lying down, and scooting sideways. Once toddlers looked away for more than two seconds, the test trial in question was aborted. Looking times to the recorded videos were re-coded for reliability. All coders, both live and video-based, were blind to the experimental condition.

To be included in the sample, children had to meet the following criteria: (1) They had to look at least five seconds to all four familiarization trials and the intertrial. (2) In test trials, inclusion required watching the screen for at least 9 seconds after the prostrating began to ensure that toddlers saw the agent submit and yield the way. If this minimum was not met, each test trial could be repeated up to three times; if still unmet, the participant was excluded. (3) Parents had to keep their eyes closed throughout the entire test phase; failure to do so resulted in the child's exclusion.

Data analysis

Statistical analyses were conducted using JASP (Version 0.19.3.0) and IBM SPSS Statistics software. The primary dependent variable was toddlers' looking time during each test trial. We used a repeated-measures ANOVA to examine the effects of trial outcome (Expected vs. Unexpected) and presentation order (Expected-first vs. Unexpected-first) on looking times.

This design allowed us to test the hypothesis that toddlers would look longer at the unexpected event—specifically, when the male-voiced agent submitted and yielded to the female-voiced agent—suggesting that they held an expectation that the male-voiced agent would prevail in the dominance conflict.

Follow-up and exploratory analyses examined the simple effects of outcome within each presentation order, as well as potential modulation by child gender or age.

Results

All reported tests are two-tailed. A repeated-measures analysis of variance (ANOVA) was conducted on looking times, with Test trial type (Expected vs. Unexpected) as a within-subjects factor and Presentation order (Expected-first vs. Unexpected-first) as a between-subjects factor. As predicted, participants looked longer when a male-voiced agent submitted and yielded to a female-voiced agent (M = 41.67, SD = 14.96) than when the male-voiced individual prevailed (M = 37.43, SD = 10.53), $F(1, 46) = 5.044$, $p = .030$, partial $\eta^2 = .10$. Specifically, looking-times to Unexpected over Expected outcomes on average were 4.3 seconds longer (see Figure 1), yielding a moderate effect (Cohen's $d = .37$).

Follow-up analyses using a MANOVA revealed no significant influence of gender or age on looking patterns, which led us to exclude these variables from the final model to improve clarity and focus on the effects of trial outcome and presentation order.

Additionally, an interaction indicated that the order in which test trials were presented influenced the looking times to Unexpected and Expected test trials, $F(1, 46) = 17.842$, $p < .001$, partial $\eta^2 = .28$, Due to this significant interaction, further analysis was performed to examine the effects of trial type within each presentation order. Toddlers who saw the unexpected trial block first spent significantly more time looking at the unexpected outcome compared to the expected one (Unexpected trials: M = 50.17, SD = 14.57, Expected trials: M = 37.96, SD = 11.02; $t(23) = 4.05$, $p < 0.001$, d = .826) (See Figure 2).

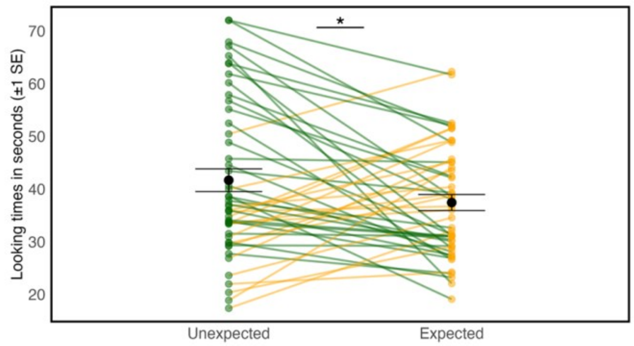

Figure 1: Looking times in seconds to unexpected versus expected test trials, with means and individual data points. Error bars indicate ± 1 SE. * = p < .05.

In contrast, children who saw the expected outcome first did not show a significant difference in looking times between the unexpected and expected trials (unexpected trials: M = 33.17, SD = 9.72, expected trials: M = 36.90, SD = 10.23; $t(23) = -1.65$, $p = .113$, $d = -.336$) (See Figure 2).

The interaction of test trial type by trial order indicates that the difference in looking times between the unexpected and expected events was primarily driven by participants who encountered the unexpected events first.

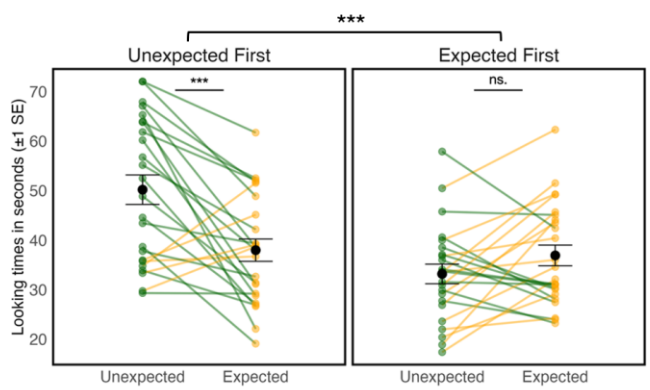

Figure 2: Interaction of test trial type and presentation order: Looking times in seconds to unexpected versus expected test trials with means and individual data points, broken down by order of test trials. Error bars indicate ± 1 SE. *** = p < .001, ns. = p > .05.

During data collection, notable qualitative observations were made. Specifically, four of the tested children verbally reacted to the agents saying "mamma" (Norwegian for *Mum*) when encountering the agent with a female voice and/or "pappa/dadda/mann" (Norwegian for *Dad* or *Man*) in response to the agent with a male voice. None made the reverse pairing.

Discussion

Using the well-validated right-of-way dominance paradigm (Thomsen et al., 2011; Thomas et al., 2018; Pun, Birch, & Baron, 2016, 2022), the current results indicated that 18–24- month-old toddlers expected male-voiced novel agents to prevail in conflict over female-voiced ones. To our knowledge, this is the first study to investigate expectations about dominance and gender at such an early age, using voice as a cue to gender. Two accounts of the observed effect are presented below.

First, the conceptualization of gender—specifically, recognizing the two agents as "male" and "female"— may have created the expectation that the male-voiced agent would prevail over the female-voiced agent. This account is consistent with studies demonstrating that by 18 months of age, children are capable of correctly matching female voices to female faces and male voices to male faces (Hillairet de Boisferon et al., 2015; Miller, Younger, & Morse, 1982, 1982; Patterson & Werker, 2002; Poulin-Dubois et al., 1994; Poulin-Dubois, Serbin, & Derbyshire, 1998; Walker-Andrews et al., 1991). Additionally, the qualitative observations we gathered (i.e., that toddlers called the female-voiced agent "mamma" and the male-voiced agent "pappa/dadda/mann") supports this account.

To further investigate whether toddlers recognize the genders of the male and female-voiced agents, future work might replicate the study reported here with an additional stimulus: a final video featuring the two agents with male and female voices, alongside a smaller third agent. Initially, the two agents would produce vocalizations (to remind the child which voice belongs to each agent) while the smaller agent observes them. The third agent would then alternates between saying "Hi Mamma" and "Hi Pappa" so that children's first looks could be analyzed, assessing whether they shift their gaze towards the female-voiced agent when "Mamma" is spoken, and towards the male-voiced agent when "Pappa" is spoken. We suggest the words "Mamma" and "Pappa" primarily because the children themselves pronounce these terms in the present study. Additionally, these are typically among the first words children learn, understand, and most closely associate with gender representation. If gaze shifts towards the male- and female-voiced characters were documented when the words "Mamma" and "Pappa" were spoken, respectively, it would indicate that toddlers associate the gendered voices with the corresponding gender concepts.

A second account of the present results, however, is that the expectation that male-voiced agents will prevail in conflict with the female-voiced agent is instead driven by the acoustic features of the voices that cue greater formidability. The two main perceptible differences between female and male voices are pitch and formants. Due to anatomical differences in the vocal tract and vocal folds, these two acoustic dimensions are relatively lower in male voices. Specifically, a deeper voice tends to be perceived as belonging to someone who is physically larger, stronger, and a better leader, and this applies to both female and male voices.

From the work demonstrating that 9–12-month-old infants use physical size as cue of social dominance (Thomsen et al., 2011), we know that infants expect larger agents to prevail in dominance interactions. Additionally, Pietraszewski et al. (2017) found that infants as young as three months are sensitive to the relationship between body size and sound frequencies, specifically expecting larger organisms to produce lower-frequency sounds. Therefore, voice pitch may influence inferences of physical size and strength, and hence function as a cue to social dominance. If so, infants should expect agents with relatively lower voices to prevail, regardless of whether it is male or female. Rather than learned gender stereotypes lending themselves to predicting the outcome of a dominance conflict, the fact that the lowest voices tend, on average, to be male, would instead lend itself to dominance predictions that may coalesce to gendered stereotypes with increasing experiences as children age.

To further investigate, a follow-up experiment might test preverbal infants, using male and female voices with manipulated fundamental and formant frequencies to create high and low frequency variants including both within- and between-gender pairings. Here children should anticipate that the agent with the lower voice will prevail in dominance conflicts, regardless of gender. Importantly, if preverbal infants expect that male-voiced novel agents will prevail over female ones—but make no association between voice and gender as indicated by their first anticipatory looks in the validation trials described above—then it would indicate that sex-typed voice characteristics, rather than nascent gender conceptualizations, license inferences that males will prevail and females should yield. Such biologically-grounded inferences may in turn develop into gendered stereotypes as children age.

Finally, future work could examine the relative role of formidability and gender concepts on predictions of dominance by directly pitting gender and other formidability cues against one another. For instance, one might test whether toddlers and infants expect that a larger female will prevail in dominance conflict over a smaller male, or vice versa.

Considerations of Competing Accounts

The present research raises further questions. Specifically, the findings could simply be driven by differential expectations about whether male- or female-voiced novel agents would fall over and scoot away, not by any predictions about dominance and submission. However, this counter-argument has been addressed in several control studies using other formidability cues.

Thomsen et al. (2011) showed that infants made no differential predictions about whether large or small novel agents would fall over and move away in the absence of zero-sum conflict (e.g., when the agents were alone on the stage, or when they moved in the same direction behind one another). Pun, Birch, and Baron (2016, 2022) further demonstrated that the violation-of-expectation effect disappears when agents are not in direct conflict (e.g., when separated by a barrier in the middle of the stage). Finally, Thomas et al. (2018) used same-sized novel agents, as in the present study, and again employed many of the same control strategies to show that toddlers selectively reached for dominant agents to whom others yield only in the context of goal conflict (but not, for example, when the puppets walk behind one another or appear alone on separate stages). While one could of course speculate that such alternative accounts might still apply when dominance is cued by voice rather than by relative size, coalition strength, or prior win–lose history, we see no a priori theoretical reason to expect these control findings to differ across modalities.

Discussion of the interaction effect: The role of test order

Crucially, we found a main effect of test type, such that toddlers looked longer at the unexpected outcome when the male agent yielded to the female one, even when controlling for the effects of presentation order. In addition to the main effect of trial type, we also found a significant interaction between trial type and test order. Specifically, toddlers looked significantly longer at the unexpected outcome when it was presented first, while the difference was less pronounced when the expected event was shown first.

Note that this pattern of results cannot be fully explained by a general decline in attention over time. If the difference in looking times were merely due to fatigue or attentional habituation, we would expect both groups to show higher attention during the first trial block and a general reduction of attention during the second. However, toddlers who saw the expected event first maintained consistent looking times across both blocks, suggesting that once they saw a male-voiced character submit to a female-voiced one, their attention did not decline. At the same time, unexpected events may be especially surprising and engaging when presented first, before an expected test trial has confirmed the default, predicted state of the world. This may more strongly reduce children's engagement and sustained attention to subsequent events that confirm their prior predictions. Overall, our results indicate that participants' responses were not solely influenced by general attentional factors but were modulated by both the content and the order of the events.

The tendency for infants and toddlers to look longer at the first test event they see, as well as the interaction between longer looking times to unexpected (vs. expected) outcomes and test order, are recurrent findings in the large body of looking-time violation-of-expectation studies (see Margoni, Surian, & Baillargeon, 2024, for a review). For instance, Thomsen et al. (2011) and Pietraszewski et al. (2017) reported similar order effects in infants' looking times. Moreover, even in the absence of a main effect, a significant interaction between trial type and test order is frequently interpreted as evidence that infants and toddlers discriminate between the two test types: If children were not sensitive to the contrast between expected and unexpected outcomes, then test order could not plausibly interact with trial type.

Conclusion

In sum, we have provided the first evidence that toddlers expect a male-voiced character to prevail in dominance conflict over a female-voiced individual, holding all other factors constant. These findings may reflect very early, learned gender stereotypes that already specify that males will dominate females. How such generative, gendered expectations could form so early—even in a relatively gender-egalitarian Scandinavian culture like Norway—remains unclear.

Alternatively, our findings could point to a mechanism grounded in the biology of male and female bodies—lower voices of men cause perceptions of greater strength and size—which itself licenses inferences of dominance and may form an early developmental origin of gender bias.

Acknowledgments

We are grateful to the participating toddlers and parents. This work was supported by the European Research Council (ERC#101040978-COORDINATE) and The Norwegian Research Council (NFR #231157/F10) (both to L.T.). F.F. was also supported by an E.U. Erasmus scholarship for student exchange and L.T. by the National Danish Research Foundation (#dnrf-144).

References

Anderson, K. J., & Leaper, C. (1998). Meta-analyses of gender effects on conversational interruption: Who, what, when, where, and how. Sex roles, 39(3), 225-252. https://doi.org/10.1023/A:1018802521676

Anderson, R. C., & Klofstad, C. A. (2012). Preference for leaders with masculine voices holds in the case of feminine leadership roles. *PloS one*, 7(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Aung, T., & Puts, D. (2020). Voice pitch: a window into the communication of social power. *Current opinion in psychology*, 33, 154-161. https://doi.org/10.1016/j.copsyc.2019.07.028

Casstevens, R. M. (2007). jHab: Java habituation software (version 1.0. 2)[computer software]. Chevy Chase, MD.

Charafeddine, R., Zambrana, I. M., Triniol, B., Mercier, H., Clément, F., Kaufmann, L., ... & Van der Henst, J. B. (2020). How preschoolers associate power with gender in male-female interactions: A cross-cultural investigation. *Sex Roles*, 83, 453-473. https://doi.org/10.1007/s11199-019-01116-x

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological science*, 25(5), 1132-1139. https://doi.org/10.1177/0956797614523297

Dovidio, J. F., Ellyson, S. L., Keating, C. F., Heltman, K., & Brown, C. E. (1988). The relationship of social power to visual displays of dominance between men and women. Journal of personality and social psychology, 54(2), 233. https://psycnet.apa.org/doi/10.1037/0022-3514.54.2.233

Esposito, A. (1979). Sex differences in children's conversation. Language and Speech, 22(3), 213-220. https://doi.org/10.1177/002383097902200302

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, 102(2), 1213-1222. http://dx.doi.org/10.1121/1.421048
Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in cognitive sciences*, 4(7), 258-267. https://doi.org/10.1016/S1364-6613(00)01494-7
Gazes, R. P., Hampton, R. R., & Lourenco, S. F. (2017). Transitive inference of social dominance by human infants. *Developmental science*, *20*(2), e12367. https://doi.org/10.1111/desc.12367
Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. Psychological bulletin, 131(6), 898. https://doi.org/10.1037/0033-2909.131.6.898
Hansen, M., Schoonover, A., Skarica, B., Harrod, T., Bahr, N., & Guise, J. M. (2019). Implicit gender bias among US resident physicians. BMC medical education, 19, 1-9. https://doi.org/10.1186/s12909-019-1818-1
Harries, M., Hawkins, S., Hacking, J., & Hughes, I. (1997). Changes in the male voice at puberty. *Archives of Disease in Childhood*, 77, 445–447. https://doi.org/10.1136/adc.77.5.445
Harries, M., Hawkins, S., Hacking, J., & Hughes, I. (1998). Changes in the male voice at puberty: vocal fold length and its relationship to the fundamental frequency of the voice. *Journal of Laryngology and Ontology*, 112, 451–454. https://doi.org/10.1017/S0022215100140757
Hillairet de Boisferon, A., Dupierrix, E., Quinn, P. C., Lœvenbruck, H., Lewkowicz, D. J., Lee, K., & Pascalis, O. (2015). Perception of multisensory gender coherence in 6- and 9-month-old infants. *Infancy*, 20(6), 661-674. https://doi.org/10.1111/infa.12088
Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical society of America*, 97(5), 3099-3111. https://doi.org/10.1121/1.411872
Keating, C. F., & Bai, D. L. (1986). Children's attributions of social dominance from facial cues. *Child development*, 1269-1276. https://doi.org/10.2307/1130449
Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. *PloS one*, 10(8), e0133779. https://doi.org/10.1371/journal.pone.0133779
Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698-2704. https://doi.org/10.1098/rspb.2012.0311
Kunst, J. R., Fischer, R., Sidanius, J., & Thomsen, L. (2017). Preferences for group dominance track and mediate the effects of macro-level social inequality and violence across societies. *Proceedings of the National Academy of Sciences*, *114*(21), 5407-5412. https://doi.org/10.1073/pnas.1616572114
Künzel, H. J. (1989). How well does average fundamental frequency correlate with speaker height and weight?. *Phonetica*, 46(1-3), 117-125. https://doi.org/10.1159/000261832
Lakoff, R. (1973). Language and woman's place. Language in society, 2(1), 45-79. https://doi.org/10.1017/S0047404500000051
Lee, I. C., Pratto, F., & Johnson, B. T. (2011). Intergroup consensus/disagreement in support of group-based hierarchy: an examination of socio-structural and psycho-cultural factors. *Psychological bulletin*, 137(6), 1029. https://doi.org/10.1037/a0025410
Leongómez, J. D., Mileva, V. R., Little, A. C., & Roberts, S. C. (2017). Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PloS one*, 12(6), e0179407. https://doi.org/10.1371/journal.pone.0179407
Lukaszewski, A. W., Simmons, Z. L., Anderson, C., & Roney, J. R. (2016). The role of physical formidability in human social status allocation. *Journal of Personality and Social Psychology*, 110(3), 385. https://doi.org/10.1037/pspi0000042
Margoni, F., Surian, L., & Baillargeon, R. (2024). The violation-of-expectation paradigm: A conceptual overview. *Psychological Review*, 131(3), 716. https://doi.org/10.1037/rev0000450
Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, *109*(18), 6862-6867. https://doi.org/10.1073/pnas.1113194109
McComb, K. E. (1991). Female choice for high roaring rates in red deer, Cervus elaphus. *Animal Behaviour*, 41(1), 79-88. https://doi.org/10.1016/S0003-3472(05)80504-4
Miller, C. L., Younger, B. A., & Morse, P. A. (1982). The categorization of male and female voices in infancy. *Infant Behavior and Development*, 5(2-4), 143-159. https://doi.org/10.1016/S0003-3472(05)80504-4
Muller, U., & Mazur, A. (1997). Facial dominance in Homo sapiens as honest signaling of male quality. *Behavioral Ecology*, 8(5), 569-579. https://doi.org/10.1093/beheco/8.5.569
Murray, G. R., & Schmitz, J. D. (2011). Caveman politics: Evolutionary leadership preferences and physical stature. *Social Science Quarterly*, 92(5), 1215-1235. https://doi.org/10.1111/j.1540-6237.2011.00815.x
Patterson, M. L., & Werker, J. F. (2002). Infants' ability to match dynamic phonetic and gender information in the face and voice. *Journal of experimental child psychology*, 81(1), 93-115. https://doi.org/10.1006/jecp.2001.2644
Pietraszewski, D., Wertz, A. E., Bryant, G. A., & Wynn, K. (2017). Three-month-old human infants use vocal cues of body size. *Proceedings of the Royal Society B: Biological Sciences*, 284(1856), 20170656. https://doi.org/10.1098/rspb.2017.0656
Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., ... & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: a meta-

analysis. *Animal Behaviour*, 95, 89-99. http://dx.doi.org/10.1016/j.anbehav.2014.06.011

Plavcan, J. M. (2001). Sexual dimorphism in primate evolution. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 116(S33), 25-53. https://doi.org/10.1002/ajpa.10011

Poulin-Dubois, D., Serbin, L. A., & Derbyshire, A. (1998). Toddlers' intermodal and verbal knowledge about gender. *Merrill-Palmer Quarterly (1982-)*, 338-354.

Poulin-Dubois, D., Serbin, L. A., Kenyon, B., & Derbyshire, A. (1994). Infants' intermodal knowledge about gender. *Developmental psychology*, 30(3), 436. https://doi.org/10.1037/0012-1649.30.3.436

Pun, A., Birch, S. A., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Sciences*, *113*(9), 2376-2381. https://doi.org/10.1073/pnas.1514879113

Pun, A., Birch, S. A., & Baron, A. S. (2022). Infants infer third-party social dominance relationships based on visual access to intergroup conflict. *Scientific Reports*, *12*(1), 18250. https://doi.org/10.1038/s41598-022-22640-z

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), 601-609. http://dx.doi.org/10.1098/rspb.2011.0829

Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and human behavior*, 27(4), 283-296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal behaviour*, 65(3), 519-530. https://doi.org/10.1006/anbe.2003.2078

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F 0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117(2), 944-955. https://doi.org/10.1121/1.1848011

Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33(5), 1208. http://dx.doi.org/10.1037/0096-1523.33.5.1208

Salles, A., Awad, M., Goldin, L., Krus, K., Lee, J. V., Schwabe, M. T., & Lai, C. K. (2019). Estimating implicit and explicit gender bias among health care professionals and surgeons. JAMA network open, 2(7), e196545-e196545. https://doi.org/10.1001/jamanetworkopen.2019.6545

Sorokowski, P., Puts, D., Johnson, J., Żółkiewicz, O., Oleszkiewicz, A., Sorokowska, A., ... & Pisanski, K. (2019). Voice of authority: Professionals lower their vocal frequencies when giving expert advice. *Journal of Nonverbal Behavior*, 43, 257-269. https://doi.org/10.1007/s10919-019-00307-0

Taylor, A. M., Reby, D., & McComb, K. (2010). Size communication in domestic dog, Canis familiaris, growls. *Animal Behaviour*, 79(1), 205-210. http://dx.doi.org/10.1111/j.1469-7998.2009.00661.x

Taylor, A. M., & Reby, D. (2010). The contribution of source–filter theory to mammal vocal communication research. Journal of Zoology, 280(3), 221-236. https://doi.org/10.1111/j.1469-7998.2009.00661.x

Terrizzi, B. F., Brey, E., Shutts, K., & Beier, J. S. (2019). Children's developing judgments about the physical manifestations of power. *Developmental psychology*, *55*(4), 793. https://psycnet.apa.org/doi/10.1037/dev0000657

Thomas, A. J., Thomsen, L., Lukowski, A. F., Abramyan, M., & Sarnecka, B. W. (2018). Toddlers prefer those who win but not when they win by force. *Nature Human Behaviour*, *2*(9), 662-669. https://doi.org/10.1038/s41562-018-0415-3

Thomsen, L. (2020). The developmental origins of social hierarchy: How infants and young children mentally represent and respond to power and status. *Current Opinion in Psychology*, *33*, 201-208. https://doi.org/10.1016/j.copsyc.2019.07.044

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. science, 331(6016), 477-480. http://dx.doi.org/10.1126/science.1199198

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, NJ: Prentice Hall.

Toscano, H., Schubert, T. W., & Sell, A. N. (2014). Judgments of dominance from the face track physical strength. *Evolutionary Psychology*, 12(1), 1-18. https://doi.org/10.1177/147470491401200101

Vannoni, E., & McElligott, A. G. (2008). Low frequency groans indicate larger and more dominant fallow deer (Dama dama) males. *PloS one*, *3*(9), e3113. https://doi.org/10.1371/journal.pone.0003113

Walker-Andrews, A. S., Bahrick, L. E., Raglioni, S. S., & Diaz, I. (1991). Infants' bimodal perception of gender. *Ecological Psychology*, *3*(2), 55-75. https://doi.org/10.1207/s15326969eco0302_1

Windhager, S., Schaefer, K., & Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *American Journal of Human Biology*, 23(6), 805-814. https://doi.org/10.1002/ajhb.21219

World Economic Forum. (2024). Global Gender Gap Report 2024. https://www.weforum.org/publications/global-gender-gap-report-2024/digest/

Zamponi, V., Mazzilli, R., Mazzilli, F., & Fantini, M. (2021). Effect of sex hormones on human voice physiology: from childhood to senescence. *Hormones*, 20(4), 691-696. https://doi.org/10.1007/s42000-021-00298-y

Zhang, J., Hodges-Simeon, C., Gaulin, S. J., & Reid, S. A. (2021). Pitch lowering enhances men's perceived aggressive intent, not fighting ability. *Evolution and*

Human Behavior, 42(1), 51-60. https://doi.org/10.1016/j.evolhumbehav.2020.07.007