Your Prosody Matters! The Effect of Controlling Tone of Voice on Listeners Throughout the Lifespan

Berdien Vrijders¹, Netta Weinstein², Silke Paulmann³, Bart Soenens¹, Joachim Waterschoot¹, Maarten Vansteenkiste¹

¹Ghent University
²University of Reading
³University of Essex

Berdien.vrijders@ugent.be, n.weinstein@reading.ac.uk, paulmann@essex.ac.uk, bart.soenens@ugent.be, joachim.waterschoot@ugent.be, maarten.vansteenkiste@ugent.be

Abstract

According to Self-Determination Theory, speakers can communicate with listeners either in more controlling or in more autonomy-supportive ways. Whereas most previous studies focused on the content of both communication styles, the current research examined whether experimentally induced controlling versus autonomy-supportive tone of voice differentially predicts listeners’ felt pressure, closeness, intention for collaboration, and feelings of fear and anger, even when listeners are exposed to these communications only briefly. In three experimental studies with adults (Study 1, \( N = 61; M_{\text{age}} = 31.51 \)), adults that are parents (Study 2, \( N = 111; M_{\text{age}} = 44.73 \)), and toddlers (Study 3, \( N = 189; M_{\text{age}} = 4.93 \)), multilevel analyses indicated that voice quality is the most critical acoustic parameter distinguishing between controlling and autonomy-supportive prosody. That is, sentences spoken with a harsher, relative to a softer, tone of voice were perceived as more pressuring, leading to higher levels of felt pressure (Study 1, 2 and 3). Listening to such harsh voices explained why listeners felt less close to and were less inclined to collaborate with controlling speakers (Study 2) and reported higher levels of anger and fear (Study 3). Results for the first time show the impact of a speaker’s tone of voice on listeners across ages, with adults and toddlers alike reporting more maladaptive effects following controlling tone of voice.

Index Terms: motivational prosody, self-determination theory, parent-child interaction

1. Introduction

Teachers asking their students to stop interrupting a classroom lecture, managers requesting a report by the end of the day, or parents asking their children to get dressed and out the door: everyday life is replete with brief but powerful interactions in which a speaker tries to direct the behavior of a listener. To this end, speakers can rely on different motivational practices. According to Self-Determination Theory (SDT; [1]), a broad theory on human motivation, these practices can vary in their level of control relative to autonomy support and this variation is consequential for listeners’ perceptions and functioning. When listeners are approached in an autonomy-supportive way, speakers empathize with and support their interlocutor’s interests, preferences, and values, which enhances their experience of free choice, volition, and self-endorsement of their actions. In contrast, when speakers motivate their listeners in a controlling way, they put pressure on them to act, think, or behave in speaker-prescribed ways and, hence, reduce listeners’ sense of free choice [1, 2].

Recent research has begun to show that, in addition to the words used by speakers (e.g., “you may” vs. “you should”; [3-5]), autonomy-supportive and controlling communication can also be differentiated on the basis of speakers’ tone of voice [6, 7]. In the present study, we examined, for the first time, whether a single sentence spoken in a controlling (as opposed to an autonomy-supportive) tone of voice (or prosody) could shape listeners’ perceptions, and whether this is true for adults as well as toddlers.

1.1. Autonomy-supportive and controlling socialization

In diverse interpersonal relationships (e.g., parent-child, teacher-student, employer-employee,...), speakers use different autonomy-supportive practices to address listeners [8-10]. Autonomy-supportive speakers provide choice to interlocutors, stimulate their initiative and offer a meaningful rationale when making requests. Moreover, they follow their interlocutors’ rhythm in performing certain tasks, and are patient as they do so [11]. In contrast, controlling speakers pressure their interlocutors to act, feel, or think in specific ways. They do so by dismissing objections, by using threats or sanctions, or by making use of subtler controlling practices such as guilt induction or love withdrawal [1, 2]. Research has revealed repeatedly, across life domains (e.g., teaching, parenting, work,...) and across age groups, that controlling communication predicts poor quality motivation [12], lower well-being [13], less long-term persistence [14], and lower performance [15] than autonomy-supportive communication (see [16, 17] for meta-analyses). Further, autonomy-supportive communication predicts a host of adaptive interpersonal outcomes, including higher attachment security [18], greater emotional reliance on others for support [19] and more voluntary disclosure of personal information between parent and child [20]. Such findings have been obtained using both cross-sectional [2], longitudinal [21], diary-based [22] and experimental methods [5, 23] and making use of both self-report measures and observations [24].

Autonomy-supportive and controlling communication styles differ not only in terms of conversation practices used by speakers (e.g., giving choice and providing a rationale versus relying on threats), but also in terms of the content or lexical-semantics of the communicated message. That is, the words speakers use can vary in terms of their level of conveyed choice relative to control [9]. Hence, the functional significance or meaning attributed to the message [25, 26] can be more...
informational or more pressuring. Specifically, autonomy-supportive speakers more often make use of inviting (e.g., “I propose”; “I ask”), suggestive (e.g., “You may...”; “You could...”) or descriptive (e.g., “I notice”) language. By contrast, controlling communication involves the use of more forceful (e.g., “You have to”) and evaluative (e.g., “Good children should do X”) language and commands (e.g., “Do this!”). Previous studies, both observational and experimental in nature, have shown that whereas controlling content impacts negatively on intrinsic motivation [4, 27] and conceptual learning [28, 29], autonomy-supportive content promotes autonomy need satisfaction [30], positive affect and motor skills learning [31] and perseverance [32].

1.2. Autonomy-supportive and controlling tone of voice

Alongside communicated content, the way in which content is delivered may also differ in its level of conveyed autonomy support relative to control. That is, the paraverbal aspects of a speaker’s message, and specifically, the tone of voice or prosody, may impact the functional significance of the message to be perceived as more controlling and pressuring or more informational and autonomy-enhancing. Tone of voice can be operationalized through different acoustic cues, including the low- or highness (pitch), sharpness or harshness as an indicator of voice quality (as measured via the distribution of energy in high-frequency energy bands) and volume (i.e., intensity or amplitude) of one’s utterances (e.g.,[33]).

A couple of previous studies using both experimental designs [7, 34] and more ecologically valid methods [35], analyzed speech patterns of speakers in terms of acoustics, and found that controlling, relative to autonomy-supportive, prosody is characterized by increased energy in higher frequency bands of the voice signal, resulting in a harsher-sounding voice. Pitch and amplitude were shown to covary with an increase in vocal energy for controlling messages, yet in different directions across studies, i.e., louder/quieter and higher/lower pitch [7, 34, 35].

Similar to the way motivational content impacts listeners’ emotional and motivational functioning, listeners have been found to respond differently to controlling and autonomy-supportive prosody. When compared with a neutral tone of voice, autonomy-supportive tone of voice led to more positive and less negative affect, increased closeness and more cooperation and effort in adolescents, whereas listening to a controlling tone of voice undermined these outcomes compared to a neutral tone condition [6, 36]. Moreover, experimentally induced controlling tone of voice elicited more pressure than autonomy-supportive tone of voice, which helped to explain why listeners reported being more likely to defy controllingly communicated messages [7]. Further, even infants, who are not able to understand verbal communications yet, were shown to attend to controlling tone of voice longer as compared to autonomy-supportive tone of voice, indicating that even before the age of one, babies are able to differentiate both types of motivational intent [37].

What remains unclear from this emerging literature on motivational prosody is whether these effects can be found for listeners of diverse ages. To this end, the current research investigated whether adults and children respond to controlling and autonomy-supportive tone of voice in the same way. It was hypothesized that controlling prosody, relative to autonomy-supportive prosody, would be perceived as more pressuring, leading to more maladaptive outcomes and this across the lifespan.

2. Method

2.1. Participants

Respectively, 61 adults (Study 1: $M_{age} = 31.51$ years, $SD_{age} = 13.62$; 51% female), 111 parents (2.02 children on average) (Study 2: $M_{age} = 44.73$ years, $SD_{age} = 9.61$; 50% female) and 189 toddlers (Study 3: $M_{age} = 4.92$ years, $SD_{age} = .79$; 50% female) participated. For Study 1 and 2, participants were recruited via Prolific Academic [38] for an online survey. For Study 3, participants were surveyed during home visits, with their parent in another room so as to not influence responses. Participants or their parents provided informed consent.

2.2. Experimental stimuli

Stimuli were developed in collaboration with two professional actors (one male, one female) with experience in improvisation theater. Actors intoned the same 24 sentences (e.g., “I don’t like you doing that”, “Let me remind you of the rules”) twice, once with an autonomy-supportive tone of voice and once with a controlling tone of voice. As such, each actor ended up intoning 48 sentences. Similar to procedures outlined in Weinstein, et al. [34], actors were informed about the nature of controlling and autonomy-supportive motivations prior to recording but were not instructed on how to intone these two different motivating styles. Instead, they were instructed to speak in a way that felt natural to them and were asked to avoid sounding angry or happy. Recordings took place in a sound attenuated room in which speakers sat at an equal distance from a high-quality microphone during recordings for each condition; actors were asked to repeat the sentence until they were satisfied with the result. Sentences were presented in randomized order to participants in semantically identical pairs, counterbalanced for type of prosody used (i.e., controlling and autonomy-supportive). Adults rated all 24 sentence pairs (Study 1 & 2), toddlers rated 6 sentence pairs (Study 3), with type of prosody as a within-subject predictor.

2.3. Measures

After each sentence, participants were asked to rate how pressuring, and supportive of choice the speaker sounded (Study 1-3), how close the listener would want to be near the speaker and to what degree s/he would want to collaborate (Study 2-3), and to which degree the listener felt sad or angry when listening to the speaker (Study 3). Adults (Study 1-2) rated sentences on a rating scale from 1 (not at all true) to 7 (absolutely true), toddlers on a rating scale from 1 (not at all true) to 4 (absolutely true). A composite perceived pressure score was created each time by subtracting the felt choice from felt pressure (i.e., the higher the score, the more pressu

3. Results

3.1. Study 1

3.1.1. Manipulation check

Table 1 provides the descriptive statistics and t-tests of between-condition effects for each acoustic parameter. Controlling for gender, significant effects were found for amplitude and voice quality, while no effect was found for pitch. This indicates that sentences were spoken more loudly as well as with increased harshness indicated through increased
high-frequency energy in the controlling condition, compared to the autonomy-supportive condition.

Table 1: Between-condition comparison in terms of acoustic parameters with means (standard deviations) and paired-sample t-tests

<table>
<thead>
<tr>
<th>Condition</th>
<th>t</th>
<th>p</th>
<th>η²_p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch (Hz)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>170.03</td>
<td>(44.1)</td>
<td>0.46</td>
</tr>
<tr>
<td>CO</td>
<td>171.35</td>
<td>(35.42)</td>
<td></td>
</tr>
<tr>
<td>Amplitude (dB)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>72.8</td>
<td>(3.61)</td>
<td>3.88</td>
</tr>
<tr>
<td>CO</td>
<td>75.04</td>
<td>(2.53)</td>
<td></td>
</tr>
<tr>
<td>Voice quality (dB)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>31.45</td>
<td>(3.97)</td>
<td>8.19</td>
</tr>
<tr>
<td>CO</td>
<td>36.73</td>
<td>(2.97)</td>
<td></td>
</tr>
</tbody>
</table>

3.1.2. Primary analyses

Unconditional models to assess intraclass correlation (ICC) showed that for felt pressure, 23% of the variance existed between, rather than within, speakers. As part of the preliminary analyses, no unique effect for the sociodemographic variables age, gender and number of children were found (all p > .05). Next, findings from the linear mixed regression model indicated that overall, sentences spoken with a controlling intent (M = 2.64) were perceived as significantly more pressuring than sentences spoken with an autonomy-supportive intent (M = -.77) (β = .55, t(2856.04) = 38.45, p < .001).

3.2. Study 2

3.2.1. Preliminary analyses

A main effect of gender was found for felt pressure (β = .07, t(99) = 2.07, p = .041), but not for collaboration or closeness, indicating that women generally perceived sentences as more pressuring than men. No other effects were found for the sociodemographic variables in prediction of the study variables (all p > .05).

3.2.2. Primary analyses

Unconditional models to assess ICC showed sufficient variability at the between-raters level for conducting full models, with an ICC of 9% for felt pressure, 27% for intent to collaborate and 25% for felt closeness. Linear mixed regression modelling showed a large effect of condition, such that sentences spoken with controlling prosody were perceived as more pressuring (β = .62, t(6133) = 67.88, p < .001), and made listeners less inclined to collaborate with (β = -.41, t(6133) = -43.74, p < .001) and be close to the speaker (β = -.46, t(6133) = -49.09, p < .001) than sentences spoken in an autonomy-supportive voice.

A multilevel mediation model was then tested, with felt pressure serving as the mediator between manipulated prosody and collaboration and closeness. As can be noticed in Figure 1, manipulated prosody related positively to felt pressure, which, in turn, related negatively to collaboration with and closeness to the speaker. The total effects between prosody and both collaboration and closeness were no longer significant when introducing felt pressure as the mediator (both p > .05), showing that controlling prosody impacts reduced collaboration and closeness by eliciting pressure.

3.3. Study 3

3.3.1. Preliminary analyses

A main effect of age was found for felt pressure (β = .16, t(145.02) = 2.85, p = .005) and anger (β = .14, t(146.24) = 2.45, p = .016), indicating that older children generally perceived sentences as more pressuring and angering than younger children. No other effects were found for the sociodemographic variables in prediction of the study variables (all p > .05).

3.3.2. Primary analyses

Unconditional models to assess intraclass correlation (ICC) showed sufficient variability at the between-raters level for conducting full models, with an ICC of 36% for perceived pressure, 33% for intended collaboration, 25% for felt closeness, 32% for anger and 34% for fear. Linear mixed regression modelling showed that sentences spoken with controlling prosody were perceived as more pressuring (β = .11, t(931.84) = 4.55, p < .001) than sentences spoken in an autonomy-supportive voice. Moreover, a controlling tone of voice evoked more feelings of anger (β = .09, t(931.55) = 3.65, p < .001) and fear (β = .10, t(928.38) = 4.20, p < .001) in toddlers. No effects were found for closeness and collaboration (both p > .05).

A multilevel mediation model was again tested, with felt pressure serving as the mediator between manipulated prosody and anger and fear. Figure 2 shows that controlling prosody related positively to felt pressure, which, in turn, related positively to feelings of anger and fear in toddlers. Felt pressure served as a partial mediator, as the direct effects between prosody and both anger and fear were not cancelled out when introducing felt pressure as the mediator. Still, results indicate that controlling prosody makes toddlers feel angry and scared by eliciting pressure.
4. Discussion

In daily life, we sometimes feel reluctant to be cooperative with a speaker and may even take physical distance from them in an attempt to show our dissatisfaction. We may even feel sad, or scared when spoken to in a controlling way. What explains our disobedience has in many cases not to do with the content of the speakers’ message, but the tone with which the message was conveyed. Despite its ecological validity, the literature on motivational prosody is still in its infancy. The present set of studies aimed to contribute to this growing body of work by examining whether variation in prosody impacts variation in felt pressure, closeness, and intended collaboration not only in adults, but already early on in the lifespan. That is, it was investigated whether children as young as 4 years old respond in the same way as adults. Two key findings deserve being highlighted.

First, a systematic effect of motivational prosody was found across three studies, with different acoustic profiles reflecting different motivational intents, leading to different responses in listeners across ages. Keeping the content of the message constant, the meaning [39] attributed to the message was found to vary as a function of how the message was expressed. That is, being spoken to in a controlling, as compared to autonomy-supportive, tone of voice made listeners feel more pressured, made them less willing to collaborate with the speaker and instead led them to take more distance from the speaker. These findings are congruent with previous research [6, 7, 36], while also extending them by showing that listeners across the lifespan (i.e., toddler to adult) pick up the variation in tone of voice as it changes from sentence to sentence. Indeed, multilevel analyses indicated that a large percentage of the variance (i.e., 73-91%) was situated at the within-person level, where the manipulation took place. Said differently, brief exposure to a controlling and autonomy-supportive tone of voice suffices to generate different experiences in both adults and toddlers, and thus, that motivational tone of voice is picked up and differentiated quickly by listeners of various ages. These findings fit well with neurophysiological research showing that adults differentiate controlling and autonomy-supportive prosody from each other and neutral tone of voice within 200 ms after sentence onset [40], and that controlling tone of voice especially is picked up early on and leads to preferential, and more in-depth processing [41].

Second, mediational analyses indicated that felt pressure explained why controlling prosody reduced participants’ intention to collaborate and undermined their felt closeness to the speaker (Study 2), as well as aroused participants’ feelings of anger and fear (Study 3). These findings are congruent with prior studies which found autonomy need satisfaction to account for the impact of controlling, relative to autonomy-supportive, language on participants’ intrinsic motivation and persistence [27, 30]. Yet, rather than focusing on intrinsically motivating activities, the requests in the current study often involved rather boring activities that one would not spontaneously engage in. The findings confirm the broader claim within Self-Determination Theory that basic need experiences have explanatory power, thereby accounting for minimal variations in the effects of the social context [42].

Finally, a few limitations need to be mentioned. First, participants’ reactions may be different in case they are not asked to imagine being exposed to the materials, but are actually exposed to a speaker addressing them with a controlling and autonomy-supportive tone of voice. Although well-controlled lab studies as the current one allow to isolate critical factors and minimize contaminating factors, a limitation is that they come with somewhat lower ecological validity. It will be important to test how participants react during real conversations and to monitor prosody effects on emotional and behavioral outcomes. Second, the current study used a within-subjects design because this approach allowed us to account for individual differences in listeners, for instance the tone of voice listeners are more habitually used to hear around them. Although the presentation of materials was balanced, hearing the contrasting tone of voice may have artificially inflated the differences in prosody. Future research might want to pair within- and between-subjects designs, which together offer a more conservative test of motivational tone of voice perception while controlling for individual differences in listeners.

5. Conclusion

In conclusion, the current set of experimental studies has shown that the tone of voice with which speakers convey their message, has an immediate impact on their interlocutors, even when listeners are as young as 4 to 6 years old. Indeed, speaking with a harsher voice makes speakers sound more controlling, which makes interlocutors feel pressured and therefore less inclined to collaborate with or be near the speaker, and more likely to experience feelings of fear or anger.

6. References

[9] J. Reeve, “Why teachers adopt a controlling motivating style toward students and how they can become more autonomy


