How parents motivate their children through prosody

Silke Paulmann¹, Berdien Vrijders¹, Netta Weinstein², Maarten Vansteenkiste³

¹University of Essex, UK
²Cardiff University, UK
³Ghent University, Belgium

Abstract

To shed light on how prosody patterns used by parents to motivate their children, the current study investigated which acoustic features define motivational speech. In particular, we focused on instances in which a controlling versus autonomy-supportive tone of voice was used by parents. To this aim, Dutch parent-child interactions were analyzed acoustically. Results of a hierarchical linear modelling analysis showed that the acoustic parameters intensity and speech rate differed significantly for both types of motivational speech. More specifically, controlling messages were uttered with a louder voice and faster rate than autonomy-supportive messages, which were conveyed using a quieter voice and slower rate. Findings support earlier research on motivational prosody and extend them to more spontaneous interactions, as well as to another language group.

Index Terms: speech processing, vocal affect, inferential communication, pragmatic language

1. Introduction

1.1. Introduction

Parents, nursery staff, teachers or sports coaches will try to energize and direct behavior of children on a daily basis. They will use so-called motivational language to help children perform certain actions (e.g., tidy the room, learn the alphabet, try to write a sentence, swim) as well as try to encourage them to keep engaging in the proposed activity. In fact, it has been argued that parent-child interactions are characterized by the frequent use of motivational utterances [1-2]. On numerous occasions will parents and children engage in motivational interactions, going from parental guidance in completing a new game [3], to involvement in children’s school achievements [4-5], and influence on children’s social behavior [6].

Obviously, parents can choose to motivate their children in different ways in these various situations. For instance, a child may need to be encouraged differently to carry out a household chore as opposed to learning how to play a new game. Theoretical underpinnings for this claim are provided by self-determination theory (SDT), which argues that two types of qualitatively different messages can drive behaviors [7-9]. First, autonomy-supportive statements such as “you may do this if you choose” suggest that the child’s own interests and beliefs are supported. They are left with a feeling of free choice of what to do and, crucially, feel as if their actions are self-endorsed [8-9]. In contrast, statements such as “you must do this” can motivate children in a more controlling way. These messages imply a sense of pressure and as such leave the child feeling with a lack of choice.

It has been repeatedly shown that controlling and autonomy-supportive language can impact differently on listeners’ performance and well-being [8-9]. Research suggests that controlling messages have more direct effects than autonomy-supportive messages [10-11], as the latter leave room for listeners to act out of their own convictions. Accordingly, when parents expect something to be done immediately (e.g., a child tidying their room), they may choose to use more controlling language (e.g., “you must tidy your room”) which, however, has been argued to undermine listeners’ well-being and intrinsic motivation in the long-term. For example, several lab-based studies exploring parent-child interactions report a decrease in the child’s intrinsic motivation, the more mothers spoke in a controlling or pressuring way [12-14; for a review, see 15].

In contrast, autonomy-supportive language can improve the relationship between conversation partners [7-9]. For instance, autonomy-supportive environments have been shown to shape higher-quality student-teacher relationships and engage students more [16], can lead to improved learning, performance and persistence in students [17], and are linked to better executive functioning in young children [18].

1.2. Nonverbal motivational language

While much focus has been put on the verbal aspects of motivational language [8-9; 12-14; 19-26], few studies have explored the supra-segmental features of motivational speech. However, this non-verbal aspect of motivational speech will be particularly important in situations where semantic cues are not motivationally biasing. For instance, “it’s time to clean your room” may be spoken in a firmer, pressure eliciting tone of voice if the parent wants the task to be carried out right away, or may be spoken in a softer, less harsh sounding tone when the parent is trying to encourage the child to see the value of doing so for themselves. Thus, it can be argued that the nonverbal aspects of motivational speech, and more specifically the prosody used when uttering motivational language, may differ between autonomy-supportive and controlling contexts.

Some past research confirms this hypothesis [27-28]. Specifically, it was shown that sentences expressing controlling contexts were intoned using a louder voice, faster speech rate, higher pitch and harsher sounding voice quality than sentences expressing autonomy-supportive statements.
Crucially, in a follow-up perception study, listeners confirmed that autonomy-supportive intonation felt less pressuring and more supportive of choice than controlling prosody.

Subsequently, the same authors explored prosody patterns for the two types of motivation in materials that lacked motivational biasing words [28]. Results mirrored and extended findings from their original study, as controlling prosody was again characterized by a faster speech rate and a louder and harsher sounding voice, while autonomy-supportive sentences were expressed with a quieter voice and slower speech rate. Interestingly, differences between these two studies were observed with regard to pitch use by speakers. Actors recruited in the second study, expressed autonomy-support with a higher mean pitch than controlling sounding sentences, while the opposite pattern had been found in the original study using non-trained student speakers.

So far, motivational prosody has only been explored in sentences addressing adult listeners, leaving the question if similar patterns can be found in adult-child interactions. Moreover, past results were all reported for English materials and other languages still require further investigation. Thus, the current study explored prosody use in autonomy-supportive and controlling contexts in Dutch parent-child interactions. Following previous research [27-28], it was expected that controlling and autonomy-supportive speech would be defined by distinct acoustic patterns as reflected in pitch, intensity and speech rate differences.

2. Methods

Spoken materials were taken from a published study [29]. Full methodological details, including the technique used to elicit and record parent-child dyadic interactions can be found in Wuyts et al. [29].

2.1. Participants

Participants were recruited from Ghentian primary schools. A total of 124 parent-child interactions were videotaped. For the current study, these video clips were initially screened for sound quality purposes. 17 clips were excluded: 11 because videos had bad sound quality and/or contained a lot of background noise and 6 because the parents in the videos were non-native speakers of Dutch. Thus, 107 video clips were acoustically analyzed.

2.2. Materials

Within the original study [29], parent-child pairs were asked to complete multiple puzzle games together (e.g., participants may have been instructed that “the princess needs to reach the prince, without passing by the wizard”). Initially, the experimenter explained the tasks to parents who subsequently were asked to explain the game to their child. This was followed by a game phase. Unrelated to the goal of the current study, parents could have been placed into one of two different instruction groups (see [29] for details). Note that the current research team was blind to the originally assigned instruction conditions.

We first converted all video clips (avi format) into audio files (.wav) using the software Audacity. Next, the second author of this paper (Dutch speaker) screened all interactions for motivational tone of voice. This was defined as any type of utterance in which the tone of voice could stimulate children to perform an action and/or praised the children for an action they had performed. Utterances were annotated independently from motivational lexical-semantic content, i.e. they were annotated merely based on tone of voice, regardless of the words used. Only utterances made by parents were of interest; children’s speech was ignored. Annotations started after the explanation phase and were carried out using Praat [30].

For the annotation process, a distinction was made between different kinds of utterances: whole sentences, phrases and single words. The label sentence was given when an utterance contained a direct verb and formed an intonational unit, even when the utterance was not fully completed. Intonational units were defined as parts of speech that are finished on an intonational level (e.g. The dragon should also go to the prince)). Utterances that did not contain a direct verb were labelled as phrases, as were utterances that did contain a direct verb but that were not completed grammatically and did not form intonational units (e.g. The dragon should also go_ to the prince). Finally, we all annotated materials were judged for their motivational quality. A distinction was made between autonomy-supportive (S) and controlling (C) motivation.

2.3. Acoustical Analysis

Following previous research [27-28], we extracted different acoustic measures for all annotated materials. These measures included mean pitch (measured in Hz), mean intensity (measured in dB), range pitch (maximum – minimum pitch as measured in Hz) and range intensity (maximum – minimum intensity as measured in dB). For 11 utterances, no data for pitch was included since this acoustic parameter appeared unmeasurable for those instances. Speech rate was also measured (in milliseconds), by dividing the total duration of speech per condition by the number of words per condition. Number of words was preferred over number of syllables, since syllables are often pronounced with great variability in spontaneous speech [31].

Two different kinds of analyses were carried out. In a first step, we conducted a discriminant analysis to confirm that selected materials contained detectable acoustic contrasts, thereby supporting the motivational quality ratings performed by the trained researcher. Next, hierarchical linear modeling (HLM) [32-34] was used to statistically analyze the acoustic data to infer how prosody patterns might differ between autonomy-supportive and controlling contexts.

2.3.1 Discriminant analyses

Four different discriminant analyses were performed: one per utterance classification (word/phrase/sentence) and one for all utterances combined. In each analysis, motivation was entered as an independent variable, with acoustic measures (mean F0, range F0, mean dB, range dB, speech rate) serving as dependent variables. Classification results revealed that overall, the model predicted 62% of autonomy-supportive and 66.1% of controlling utterances correctly. Results also revealed that the model predicted autonomy-supportive category membership more accurately when looking at words (67.4%) as opposed to phrases (64.8%) and sentences (60%). For controlling category membership, the model classified phrases (64.9%) more accurately than words (64%) and sentences (63.3%), although accuracy rates did not differ much. Crucially, these results suggest that materials indeed...
contained discernible acoustic features that helped distinguish between the two different types of motivation.

2.3.2 Hierarchical Linear Modeling

We used HLM to statistically analyze a data structure where individual sentences (defined at level 1) were nested within speakers (defined at level 2). This method was used since it accounts for shared variance, which stems from the interdependence of data collected from the same participant as well as variation between participants and conditions [27;34]. Intraclass correlation (ICC) was calculated to assess the amount of within-subjects (i.e. between-utterances) and between-subjects variability at all levels of the hierarchy. Since variance existed at level 1 and 2, predictor variables were entered at different levels.

Full models predicted the relationship between acoustic parameters mean F0, range F0, mean dB, range dB and speech rate per word (level 1 outcome variable) and the motivation condition (level 1 predictor variable). As level 2 predictor variable, we also included the assigned instruction condition from [29]. Interactions between predictors were also predicted. As the level 2 predictor variable was only included to confirm that the original assigned instructions did not impact on motivational prosody use, only significant effects will be reported for this variable. Sentence part was entered as a level 1 control variable, to ensure that results could not be influenced by a difference in audio sample length. Hierarchical linear models were performed for each acoustic feature separately. Level 1 variables were centered around the group mean, while level 2 predictors were centered around the grand mean.

Table 1 presents means for all acoustic measures of interest for autonomy-supportive and controlling motivation separately. For mean pitch, mean intensity and speech rate, the table also shows the range within which each of these measures fell.

Table 1: Acoustic parameter means.

<table>
<thead>
<tr>
<th></th>
<th>Autonomy-Support</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean pitch (Hz)</td>
<td>210 (80-441)</td>
<td>201 (91-415)</td>
</tr>
<tr>
<td>pitch range (Hz)</td>
<td>139 (40 – 63)</td>
<td>119 (41 – 75)</td>
</tr>
<tr>
<td>mean loudness (dB)</td>
<td>56 (90-1000)</td>
<td>58 (120 – 850)</td>
</tr>
<tr>
<td>loudness range (Hz)</td>
<td>22 (40-1000)</td>
<td>24 (120 – 850)</td>
</tr>
<tr>
<td>speech rate (ms)</td>
<td>330 (91-415)</td>
<td>300 (120 – 850)</td>
</tr>
</tbody>
</table>

Pitch measurements

No differences were observed in mean pitch use between utterances spoken in a controlling versus autonomy-supportive prosody, b = -1.64, t(1279) = -.63, p = .57.

For pitch range measurements, results showed that autonomy-supportive prosody was expressed with a significantly greater pitch range than utterances deemed to sound controlling, b = -12.05, t(1279) = -2.25, p = .02, suggesting that controlling messages were spoken with a more constant pitch than autonomy-supportive messages.

No effects involving the assigned instruction predictor variable were observed, confirming that original experimental assignments did not influence prosody use in parents.

Intensity measurements

Results showed that samples from the controlling condition were spoken louder than those from the autonomy-supportive condition, b = 1.3, t(1287) = 4.84, p < .001. In addition, there was a significant interaction between the originally assigned instruction condition and the motivation condition, b = 1.14, t(1287) = 2.13, p = .03. Specifically, it was shown that parents uttering controlling materials spoke louder if they were also originally assigned to controlling game instructions as compared to parents uttering controlling prosody but originally assigned to autonomy-supportive game instructions.

For range in amplitude, results showed that parents using controlling prosody used a wider amplitude range than those expressing autonomy-supportive prosody, b = .77, t(1287) = 2.16, p = .03, indicating that the loudness with which controlling messages were uttered varied more than the loudness with which autonomy-supportive messages were spoken. No other significant effects were found.

Speech rate

Results showed that controlling sounding materials were spoken faster than those rated to sound autonomy-supportive, b = -.01, t(1287) = -2.54, p = .01.

3. Discussion

The current study set out to explore whether a distinct prosody pattern can be described for parents using controlling and autonomy-supportive motivational language. Previous research [27-28] suggests that native speakers of English expressing motivational prosody (in lab settings) modulate pitch, intensity, and speech rate differently depending on whether they aim to sound pressuring or supportive. Here, we aimed to extend this research to a different language (Dutch) and to explore whether similar effects could be found in non-scripted parent-child interactions.

In line with past reports [27-28], the present findings revealed significant differences in pitch, intensity and speech rate measurements for parents judged to sound autonomy-supportive versus controlling. While the judgements of our materials were only carried out by one researcher, additional support for the validity of the classification for materials was provided by a discriminant analysis. This showed that motivational quality predictions based on several acoustic parameters was above chance-level and similarly successful for autonomy-supportive and controlling utterances (62% vs. 66%). Supplementary discriminant analyses run separately for different sentence parts, revealed that autonomy-supportive motivation was better predicted if prosody was confined to single words as opposed to longer phrases or sentences. This result might suggest that supportive prosody is uttered with a louder and faster voice than autonomy-supportive materials. Moreover, parents using autonomy-supportive prosody restricted their use of loudness more than parents uttering controlling prosody. These results thus confirm earlier findings by Weinstein and colleagues [27-
and extend them to a different language and a different situational context (lab based vs non-scripted materials). In addition to exploring acoustic correlates of motivational prosody used by parents when trying to motivate children to do a puzzle, the current data set also allowed us to explore if the tone of voice used by parents could affect the children’s overall motivation states (data taken from [29]). Results from correlation analyses showed that the louder parents talked, the more controlled children felt. These results are in line with results reported by Weinstein et al. [27] which reports that adult’s overall well-being measures decreased after listing to controlling tones of voices for approximately two minutes. This also once more shows that intensity seems to be one of the acoustic parameters that is critical when communicating controlling motivation.

Interestingly, findings from the HLM did not reveal any significant differences for mean pitch use in parents choosing an autonomy-supportive versus controlling tone of voice. As such, these results are in conflict with reports revealing mean pitch use differences between these two motivational types [27-28]. However, while both previous studies in English showed mean pitch differences between speakers expressing controlling versus autonomy-supportive prosody, results for each study revealed that pitch was modulated in different directions. Similarly, the direction of effects for pitch range differs between the present study and Weinstein et al. [27]. Specifically, here we showed that controlling prosody was expressed with a smaller pitch range than autonomy-supportive prosody, the opposite was reported by Weinstein et al. [27]. Moreover, another study [28] failed to find any pitch range differences between the two motivational qualities. Given these diverging results, it is possible that the acoustic parameter pitch is not a significant contributor in determining perceived control or autonomy-support. That is, although pitch is an important prosodic factor for motivational utterances, it seems as if any modulation in pitch can be perceived as autonomy-support or control if it is heard alongside specific intensity, voice quality and speech rate parameters. This hypothesis fits well with reports on the perception of emotional prosody [35-36]. These previous findings, for instance, showed that similar to control and autonomy-support, anger and happiness can often display high pitch variation between studies. Furthermore, manipulation of pitch does not seem to affect the perception of either of these two emotions.

Additional support for the claim that pitch is not the driving force to convey different motivation types is offered by a perception study reported in [28]. There, it was shown that pitch was the only acoustic parameter investigated that could not predict how controlled or supported listeners felt. Moreover, it was observed in this study that pitch could be varied by different speakers – an observation which allows the assumption that pitch may be modulated even more by different speakers in real-life interactions. In turn, this assumption would explain the results that were found and uphold the hypothesis that mean pitch as measured across longer segments is not the principal factor for communicating motivational tone of voice.

Finally, it should be noted that it has been argued previously [e.g., 28] that materials elicited in the lab, may lead to more stereo-type confirming expressions. In other words, lab participants may choose to use more expressive prosody than those in non-scripted interactions. Specifically, speakers in the lab are often encouraged to sound as convincing as possible; here, however, materials were taken from a previous investigation that never aimed to explore prosody patterns. However, parents were aware that they were taped during the puzzle game which may have prompted some of them to be less explicitly controlling towards their children than they would have been otherwise. Future studies will have to explore this possibility in more detail.

4. Conclusion

In sum, the findings reported here support the hypothesis that distinct prosody patterns are used when motivating others in an autonomy-supportive or controlling way. They also extend past results to a new language and context, suggesting that motivational language is expressed similarly across languages, comparable to what has been found for emotional prosody [36]. Finally, results suggest that the acoustic parameters intensity and speech rate are crucial when conveying these two distinct types of motivation. Specifically, controlling motivation is characterized by a louder voice and a faster speech rate than autonomy-supportive motivation. Pitch, on the other hand, does not seem to be a driving force.

5. References


