Differences between the acoustic typology of autonomy-supportive and controlling sentences

Netta Weinstein, Konstantina Zougkou and Silke Paulmann

Department of Psychology & Centre for Brain Science, University of Essex, UK

`netta@essex.ac.uk, kzougo@essex.ac.uk, paulmann@essex.ac.uk`

Abstract

The current study was first to describe distinct patterns of prosody that discriminate motivationally laden speech. To do this we applied self-determination theory, a widely used motivational framework. Participants in the US and UK were asked to read out loud either autonomy-supportive sentences (that support choice and volition) or controlling (pressuring and coercive) sentences. Data analyses were conducted using a conservative hierarchical linear modeling approach to account for nesting of sentences within individuals. Across both countries and controlling for gender, autonomy-supportive sentences were read using lower pitch, less intensity, and a slower speech rate than were controlling sentences. Multiple regression analyses showed links between these patterns of prosody for each participant and his or her current level of motivation, providing additional validity to results. Findings inform both the motivation and prosody literatures and offer a first description of how different kinds of motivational speech may sound.

Index Terms: autonomy-support, controlled motivation, motivational prosody, social prosody, prosodic contour, sentence prosody

1. Introduction

Motivational language is a significant aspect of daily interactions across relationships; it is a fundamental component of people's influence on, and responses to, others. People regularly attempt to shape the behavior of others with motivationally laden language. A child may be told to "go and tidy your room" and while the verbal message plays an important role in shaping behavior, there is reason to believe that prosody is also critical; for example, listening to a calm but firm tone of voice saying "go and tidy your room now" will have greater impact on children's behavior than listening to the same words spoken in a shaky, low tone of voice.

Human communications are often intended to drive and shape others' behaviors, and in these cases motivational elements are imbued in language. The theoretical framework provided by self-determination theory (SDT) argues that two types of qualitatively different messages can drive behavior. Autonomy-supportive statements, such as "you may [do this] if you choose" improve relationships and well-being by supporting perceived choice and volition in the listener. In contrast, controlling statements such as "you must [do this]" are experienced as pressuring or coercive and can undermine well-being [1], but they may also be more effective for achieving desired behaviors in the short-term. The motivational qualities of an interaction have numerous impacts on the personal, relational, and behavioral outcomes of that interaction. Autonomy-supportive versus controlling environments have been shown to shape defiant behaviors [2], encourage interpersonal closeness [3], increase well-being [4], lead to more responsible decision-making [5], and shape

successful or unsuccessful performance on important tasks [6], among other outcomes. Furthermore, these motivational communications are relevant in political communications, parent-child relationships, sports and exercise, education, the workplace, and healthcare [7], and as such this fundamental work on motivational communications can be extended across the gamut of human interactions.

Although social psychology has a long tradition of exploring how individuals deliver and understand motivationally laden ideas expressed through specific verbal messages (i.e. words), and research emerging from SDT has examined the words used in both types of messages [8], there is no research to date on how prosody distinguishes these qualities of motivation. Attention to vocal attributes of utterances is crucial in everyday life, particularly in situations when semantic cues are missing (e.g. "time to leave", can be said in either a controlling or autonomy supportive way) or when semantics and prosody mismatch. However, the basic semantic analyses commonly used do not capture the important effect prosody has on motivational language processing. Thus, the current study set out to explore the role prosody plays when communicating motivational language. Past research on another function of prosody, namely emotional prosody, has highlighted that the interplay of multiple acoustic parameters such as frequency variables (e.g., mean, range, contour of pitch), voice qualities (e.g., jitter, shimmer, spectral frequencies), intensity (loudness), and speaking rate, are reliably associated with specific emotional states. For instance, angry statements are often expressed with high intensity, high pitch, and fast speaking rate, while sad statements are conveyed with reduced intensity, low pitch, and slow speaking rate [9]. Thus, in spoken communication, listeners can rely on differences in temporal, pitch, and intensity cues to infer how someone feels; presumably, the same is true when individuals hear motivational messages. It is crucial to note that although basic emotions and motivational styles may correlate, for example, someone who is angry may use a controlling tone, the two constructs are conceptually and operationally distinct – someone who is angry may still use autonomy-supportive language. Accordingly, emotions and motivation should show related but distinct prosody patterns, so that understanding basic emotions does not translate to defining motivational prosody patterns. Specifically, we expected that control and autonomous motivational sentences would be defined by distinct acoustic configuration profiles, as reflected in pitch, temporal, and amplitude differences for the two sentence types.

2. Method

2.1. Participants and procedures

Participants were 100 students recruited from a small university in the North-East USA ($n = 51$) and from a

comparably sized university in the East of England (*n* = 49). Of these, 39 American and 31 British females took part. Participants were aged 18 to 32 (*M* = 22 years).

Participants were randomly assigned to an Autonomy-Supportive or Controlling condition (2-way between-subjects design), which determined the motivational content of the sentences to be read out loud. Instructions for this part of the study asked participants to: "...*read a number of sentences out loud... Please read them in a loud and expressive voice, and pronounce clearly. As if you really mean it...*" Sentences were then presented one at a time at the center of a computer screen at a rate of 12 seconds per sentence. Participants first practiced on a set of 10 motivationally neutral sentences, including "join me at the park" and "why don't we meet tomorrow?" Following this, participants were asked to say aloud the sentences specific to their allocated condition multiple times. The first reading of motivational sentences was designed to increase participant immersion in the particular motivational state, though we have focused on data collected during a second reading.

2.2. Materials

2.2.1. Autonomy supportive and controlling sentences

A series of 12 each autonomy-supportive and controlling sentences were selected based on theoretical considerations and previous research in SDT [10]. Examples of these are "you may do this if you choose" and "you're free to do this" (autonomy) and "you have to do it my way" and "you ought to do it" (control). Conditions were matched on the number of words – both sets of sentences ranged from 3 to 9 words.

To describe the acoustic typology of autonomy and controlling speech, the most commonly studied acoustic parameters, pitch and intensity (mean, minimum, maximum and range), were measured in the current study. Speech rate was also measured that reflected duration of reading out a sentence (in sec.) divided by the number of syllables in that sentence (sentences ranged from 3-10 syllables each).

A rating study was conducted to validate sentences. In particular, we aimed to confirm that autonomous sentences were perceived by others to be more supportive of choice, and that controlling ones were perceived by others to be more pressuring. To this end, an independent sample (*n* = 33) of British participants listened to recordings of both controlling and autonomy-supportive sentences. Sentences were presented in randomized blocks grouped by condition, and order of sentences was randomized within each block (within-subjects design). Following each sentence, participants were asked to report on the extent each sentence was pressuring (for the controlling condition) using a scale of 1 (*not pressuring at all*) to 5 (*very pressuring*) and on the extent speakers provided choice (for the autonomous condition) with a scale of 1 (*does not support choice*) to 5 (*supports choice very much*). Results showed that participants perceived controlling sentences to be more pressuring (*M* = 3.80) than supportive sentences (*M* = 2.16), *F*(1, 32) = 133.57, *p* < .001, and less supportive of choice (*M* = 2.06) than autonomy supportive sentences (*M* = 3.62), *F*(1, 32) = 31.24, *p* < .001.

Participants also reported on their current levels of controlled motivation after each block of sentences (blocks presented multiple sentences from one condition only). Findings showed exposure to multiple sentences altered motivational states in a way consistent with the framing of the sentence. After hearing controlling sentences participants reported more controlled motivation (*M* = 3.14) as compared to after hearing autonomy supportive sentences (*M* = 2.72), *F*(1, 32) = 13.73, *p* = .001.

2.2.2. State motivation

Participants taking part in the main study, i.e. participants who read aloud the different sentences, reported on their state levels of controlled motivation after a prompt delivered at the end of the study: "how much do you feel this way right now?" Three items assessed motivation, namely feeling "pressured", "coerced", and "choiceless". Each was paired with a seven-point scale ranging from 1(not at all) to 7(very much). Internal reliability was acceptable, α = .81.

3.   Results

3.1. Analytic strategy

Primary analyses were conducted with hierarchical linear modeling (HLM) [11], [12] because individual sentences (defined at level 1) were nested within speakers (defined at level 2). This method recognizes interdependence of sentences collected from the same participant as well as variation between participants and condition. We first conducted unconditional models to assess intraclass correlation (ICC); this analysis provided an estimate of the variability within-speakers (between sentences) and between-speakers and all parameters showed sufficient variability at both levels for conducting full models. Full models predicted major prosody parameters from gender, country of origin, and condition, as well as the interactions between predictors. The order in which we entered variables was determined by conceptual considerations. Level 2 variables were centered on sample means as recommended by Bryk and Raudenbush [11]; no predictors were specified at level 1. Degrees of freedom for these models are based on the number of participants and observations included in specific analyses, accounting for limited missing data (<5%) due to problems with recordings.

3.2. Preliminary findings

3.2.1. Manipulation check

Preliminary univariate analyses of variances (ANOVAs) predicted state levels of controlled motivation from assignment to condition, gender, country, and their interactions. Findings showed no effect of sex or country on state motivation, *F*(1, 87) = .14, *p* = .71, and *F*(1, 87) = 1.01, *p* = .32. However, those in the Autonomy condition reported less controlled motivation (*M* = 1.80) at the end of the study as compared to those in the Controlling condition (*M* = 2.49), *F*(1, 87) = 6.26, *p* = .02. There were no interactions between condition, gender, and country, *F*(1, 87) < 2.80, *ps* > .10. This finding confirms the effectiveness of the sentences used in the two conditions as a mean of motivation manipulation, indicating that assignment to condition did effectively shape speakers' motivational profiles.

3.2.2. Mean levels within countries and genders

Table 1 presents findings for mean scores in each of the two countries and for each gender. Since each of these two predictors may be expected to impact the effects of condition

on prosody indicators, both were controlled for in later analyses.

Table 1. *Means for primary variables of interest in both countries tested and for each gender.*

	M USA	*M* UK	*M* Men	*M* Women
Pitch mean	202.50	186.43	148.39	214.38
Pitch min	143.14	143.15	103.12	160.98
Pitch max	336.76	293.66	261.11	338.02
Pitch range	193.62	150.51	158.66	177.04
Intensity mean	69.35	56.62	63.30	62.83
Intensity min	54.34	32.20	40.26	44.47
Intensity max	82.22	70.09	76.45	76.08
Intensity range	27.88	37.89	36.19	31.61
Speech rate	0.20	0.23	0.21	0.22

3.3. Primary findings

Separate hierarchical linear models predicted mean, minimum, maximum, and range pitch and intensity, as well as speech rate per syllable from gender, country, and condition, and from their interactions.

3.3.1. Pitch measurements

Results showed women spoke in a higher pitch than men, $b = 64.88$, $t(91) = 8.73$, $p < .001$, though there was no difference across countries, $b = 7.34$, $t(91) = 1.20$, $p = .23$. Controlling for the two covariates, condition marginally affected pitch, $b = -10.64$, $t(91) = -1.90$, $p = .06$, such that those in the Autonomy condition used lower pitch than those in the Control condition (see Table 2 for a summary of findings and effect sizes). Neither gender, $b = 14.73$, $t(89) = 1.10$, $p = .28$, nor country, $b = 14.52$, $t(89) = 1.22$, $p = .23$, interacted with condition, indicating motivation was expressed in similar ways across subsamples. In separate models, it was found that autonomy-supportive sentences had marginally lower minimum pitch, $b = -6.79$, $t(91) = -1.86$, $p = .06$, while results for maximum pitch did not show a significant difference, $b = -22.61$, $t(91) = -1.49$, $p = .14$; as well, range did not differ across conditions, $b = -15.81$, $t(91) = -1.08$, $p = .28$.

3.3.2. Intensity measurements

Generally, women spoke with lower intensity than did men, $b = -2.01$, $t(91) = -2.14$, $p = .04$, as did participants from the UK versus the US, $b = 13.02$, $t(91) = 15.84$, $p < .01$. Controlling for these gender and country differences, participants in the Control condition spoke with a louder tone of voice than those in the Autonomy condition, $b = -2.59$, $t(91) = -3.29$, $p = .002$. Condition did not interact with either gender, $b = -1.52$, $t(89) = -0.87$, $p = .39$, or country, $b = -0.31$, $t(89) = -0.21$, $p = .83$. It was also found that controlling sentences were spoken with a higher maximum intensity, $b = -2.49$, $t(91) = -2.48$, $p = .02$, though there was no effect of condition on minimum intensity, $b = -0.20$, $t(91) = -0.20$, $p = .84$. As well, autonomous sentences had marginally less range in intensity, $b = -2.29$, $t(91) = -1.98$, $p = .05$, controlling for country, $b = -9.12$, $t(91) = -8.28$, $p < .001$, and gender, $b = -3.83$, $t(91) = -2.79$, $p = .007$. Condition did not interact with either country, $b = 0.91$,

$t(91) = 0.47$, $p = .64$, or gender, $b = -2.34$, $t(91) = -0.93$, $p = .36$, in predicting range in intensity.

3.3.3. Speech rate

Results from full HLM models indicated participants from the US spoke at a faster rate, $b = -0.02$, $t(92) = -2.83$, $p = .006$, though no gender differences were apparent, $b = 0.01$, $t(91) = 0.93$, $p = .36$. Autonomous sentences were spoken at a slower rate, $b = 0.03$, $t(92) = 4.79$, $p < .001$.

Table 2. *Effects of condition for all acoustic indicators: Results from primary HLM models.*

	b	*t*	*d*
Pitch mean	-10.64	-1.90♦	.40
Pitch min	-6.79	-1.86♦	.39
Pitch max	-22.61	-1.49	.31
Pitch range	-15.81	-1.08	.23
Intensity mean	-2.59	-3.29**	.69
Intensity min	-0.20	-0.20	.04
Intensity max	-2.49	-2.48*	.52
Intensity range	-2.29	-1.98*	.42
Speech rate	0.01	1.96*	.41

Note: effect size was computed using formulation for Cohen's *d*. ♦marginal significance, **p* ≤ .05, ***p* < .01.

3.4. State level outcomes of motivational sentences

To link prosody to state levels of motivation at the end of the study, a multiple linear regression analysis was used regressing the reported state levels of controlled motivation from speakers' tone of voice (prosodic profile) in the controlling condition, accounting for covariates. Mean pitch, intensity, and speech rate (reversed) were standardized and averaged to construct a controlling prosody profile; higher scores reflected more expression of a controlling prosodic tone of voice. Results showed use of controlling tones was linked to a more controlling state orientation at the end of the study, $β = .24$, $t(91) = 2.51$, $p = .01$; furthermore, in a second step of the analysis, this effect did not interact with covariates, $βs < +/-.28$, $t(89) < 0.59$, $p > .16$. These findings indicated that those who used a more controlling profile were then more likely to report a motivational style consistent with their tone, providing additional support that these profiles reflected an expression of one's motivational state.

3.5. Supporting analyses with exemplar files

Two independent research assistants blind to hypotheses and previous findings were asked to identify audiofiles that sounded natural, as if spoken and not read, and determined that twenty participants spoke in that ideally natural tone of voice. Sentences from this subset of participants were subjected to analyses of variances (ANOVAs) to attempt to replicate findings from the overall sample presented above. As before, analyses controlled for gender and country; findings for condition are presented below.

3.5.1. Pitch

Results of ANOVAs showed that autonomous sentences (*M* = 178.7) were spoken with lower pitch, $F(1, 246) = 5.52$, $p =$

.02, than controlled sentences ($M = 188.3$). Along with a lower mean, autonomous sentences showed lower minimum pitch, $F(1, 246) = 9.97$, $p = .002$ ($M_{autonomy} = 124.2$ vs. $M_{control} = 135.2$), as well as a lower maximum pitch, $F(1, 246) = 4.32$, $p = .04$ ($M_{autonomy} = 263.8$ vs. $M_{control} = 297.6$). As was the case for full HLM models, condition did not predict range in pitch, $F(1, 246) = 2.02$, $p = .16$ ($M_{autonomy} = 139.7$ vs. $M_{control} = 162.4$).

3.5.2. *Intensity*

Similarly to pitch, autonomous sentences ($M = 62.1$) were expressed using lower intensity, $F(1, 246) = 5.69$, $p = .02$, than controlled sentences ($M = 63.2$). Sentences showed higher minimum intensity in the Autonomy condition, $F(1, 246) = 8.55$, $p = .004$ ($M_{autonomy} = 44.1$ vs. $M_{control} = 41.0$), and no differences in maximum intensity, $F(1, 246) = 0.90$, $p = .35$ ($M_{autonomy} = 77.5$ vs. $M_{control} = 78.2$). In a separate analysis, autonomous sentences had a lower range of intensity, $F(1, 246) = 8.73$, $p = .003$ ($M_{autonomy} = 33.4$ vs. $M_{control} = 37.2$).

3.5.3. *Speech rate*

Finally, autonomous sentences also showed a slower speech rate than did controlled sentences, $F(1, 246) = 7.72$, $p = .006$ ($M_{autonomy} = .016$ vs. $M_{control} = 0.18$).

4. Discussion

The present study set out to explore the acoustic-perceptual underpinnings of motivational prosody using an experimental design. In contrast to studies investigating the acoustic profiles underlying emotional prosody (e.g. [9]), we used a large number of untrained speakers to develop an initial acoustic typology of motivational speech. Findings showed that autonomy-supportive messages such as "you're free to do this" were spoken with a lower mean pitch, lower mean intensity and were read more slowly than controlling sentences. These messages were also expressed using a smaller pitch range than controlling sentences. In contrast, latter messages such as "you ought to do it" were expressed with a higher maximum intensity (i.e. louder) than autonomy supportive sentences. The observed differences in pitch, amplitude, and temporal characteristics of our motivationally laden utterances suggest that speakers adopt specific prosodic speech patterns when communicating autonomy-supportive and controlling motivational sentences.

Everyday motivationally laden messages are used to shape listeners' behaviors. For instance, doctors encourage patients to adhere to preferred health and medical procedures, teachers aim to educate and socialize their students, and parents try to energize their children to engage in valued and avoid harmful behaviors. Even in non-vertical relationships between friends, romantic partners, or housemates, motivational messages are shared daily for important tasks (e.g. financial or lifestyle decisions that affect both partners) and small tasks (e.g. cleaning one's room). Previous work has examined which words are used to discriminate different motivations [8], but next to nothing is known about motivational communication beyond words, in the tone of voice used to express these messages. Our findings support the view that perceptually distinct vocal profiles are used to express motivational speech. Moreover, we show that listeners actually perceive auditorially presented autonomous sentences as being more autonomy-supportive or supportive of choice,

whereas sentences from the controlling condition were perceived to be more pressuring. This suggests that expressive tones of voice can be formed and used to drive others to action, at least when prosodic and lexical-semantic cues are used concurrently.

Three major acoustic parameters were measured to identify vocal indicators of motivation: pitch, intensity, and speech rate. Interpretations about the direction of effects have to remain speculative at this point, but, arguably, controlling speech requires more 'effort' than autonomy controlled speech given that one is explicitly trying to shape someone else's behavior. Indeed, we find that the amount of energy used to produce controlling sentences is higher than when producing autonomy-supportive messages. Not only is this extra emphasis missing in the latter condition, autonomy-supportive speech is also characterized by slower speech rate, suggesting that the speaker might be less pressured and more flexible in their attempts to influence others' behaviors. In fact, it is well known that physiological modifications of the systems involved in speech production (i.e. respiratory, phonatory, articulatory) systematically alter a speaker's tone, and research in SDT has linked control to more physiological stress [13].

Our aim is that this research is seen as a starting point to outline different prosody patterns as a function of the speaker's motivational state and intention. Future studies can elaborate on this research in a number of ways. For example, this sample was relatively young and sentences read in a tightly controlled lab setting. This research should be replicated with older participants and those imagining themselves in real-life interactions. Furthermore, we tested a number of indicators of prosody, but future studies could further explore the exact prosody typology that characterizes autonomous and controlled motivation by examining a wider range of acoustic parameters. Finally, future studies will have to confirm acoustic profiles for the two different motivational states, independent of the content used to express a message. One way to do this is to use semantically neutral sentences intoned in either motivationally laden state.

5. Conclusions

The current study was aimed at classifying motivational speech patterns to develop an acoustic typology for autonomous and controlled motivations; ultimately, our goal is to advance a richer understanding of the nature of motivational communication. This is important for a number of reasons. First, prosody research has yet to be generalized to the large body of motivational research in SDT that manipulates motivation – the very manipulations used in the field can be modified and constructed systematically to reproduce expressive tones of voice that affect motivation and subsequent behaviors. Second, this research can be applied in analyzing real life recordings when individuals interact, to examine how motivation functions in daily life. Finally, speech inconsistencies and deceptive motivational influences can be studied for the first time once tone of voice is understood and defined as independent from words. More importantly, this research is fundamental to understanding the basic nature of autonomous and controlled motivations and how these can be communicated in tone of voice.

6. References

[1] Deci, E. L. and Ryan, R. M., "The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior", Psychol. Inq., 11:227-268, 2000.

[2] Vansteenkiste, M., Soens, B., Van Petegem, S. V. and Duriez, B., "Longitudinal associations between adolescent perceived degree and style of parental prohibition and internalization and defiance", Dev. Psychol., in press.

[3] Vansteenkiste, M., Ryan, R. M. and Deci, E.L., "Self-determination theory and the explanatory role of psychological needs in human well-being" in L. Bruni, F. Comim, & M. Pugno [Eds], Capabilities and Happiness, 187-223, Oxford: Oxford University Press, 2008.

[4] Deci, E. L., La Guardia, J. G., Moller, A. C., Scheiner, M. J. and Ryan, R. M., "On the benefits of giving as well as receiving autonomy support: Mutuality in close friendships", Pers. Soc. Psychol. B., 32:313-327, 2006.

[5] Huang, K-L., "The effect of different kind of risk-behaviors on driving and decision-making behavior using SDT", Thesis, Yun-Tech, 2006.

[6] Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M. and Deci, E. L., "Motivating learning, performance, and persistence: The synergistic role of intrinsic goals and autnomy-support", J. Pers. Soc. Psychol., 87:246-260, 2004.

[7] Weinstein, N., "Human motivation and interpersonal relationships: Theory, research, and applications", Dordrecht, NE:Springer, 2013.

[8] Ryan, R. M. and Deci, E. L., "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being", Am. Psychol., 55(1):68-78, 2000.

[9] Banse, R. and Scherer, K.R., "Acoustic profiles in vocal emotion expression", J. Pers. Soc. Psychol., 70(3):614-636, 1996.

[10] Deci, E. L. and Ryan, R.M., "Facilitating optimal motivation and psychological well-being across life's domains", Can. Psychol., 49:14-23, 2008.

[11] Bryk, A.S. and Raudenbush, S.W., "Hierarchical Linear Models in Social and Behavioral Research: Applications and Data Analysis Methods", [First Edition], Newbury Park, CA: Sage Publications, 1992.

[12] Raudenbush, S.W. and Bryk, A.S., "Hierarchical Linear Models" [Second Edition], Thousand Oaks: Sage Publications, 2002.

[13] Weinstein, N. and Ryan, R. M., "A motivational approach to stress response and adaptation", Stress & Health, 1: 4-17, 2012.