

THE EFFECTS OF DIFFERENT LEVELS OF AMPLITUDE VARIATION ON PERCEIVED SPEAKER DOMINANCE

Sara Pearsell¹, Daniel Pape¹

¹ McMaster University, Hamilton, Canada
pearsels@mcmaster.ca, paped@mcmaster.ca

ABSTRACT

The purpose of this study is to examine how suprasegmental amplitude manipulation on different linguistic levels (sentence, word, syllable) would affect perceptual dominance ratings. Six (Canadian) English speakers produced sentences/paragraphs with neutral valence. Amplitudes were then manipulated on three levels: (i) *sentence/paragraph* (i.e. amplitude enhancement of entire phrases or paragraphs), (ii) *word* (one specific word in focus, realized with enhanced amplitude), and (iii) *syllable* (i.e. increasing/decreasing amplitude differences between all stressed versus unstressed syllables). 71 participants of an online perception experiment then rated the stimuli on 4 continuous sliding scales corresponding to concepts of social and physical dominance. Results showed that amplitude increases on *sentence* level significantly increased dominance ratings. In contrast, on the *syllable* level significantly decreased dominance ratings were observed for increased amplitude levels. *Word* level manipulations and the decreased syllable condition did not have an effect on ratings.

Keywords: amplitude variation, perceived personality traits, speech perception, variability in speech

1. INTRODUCTION

In more recent research ([1], [2], [3] [4]), there has been an increase in examining the acoustic properties that influence the perception of speakers' personality traits. Research has identified influences of pitch ([5], [6], [1], [2]), pause-filled-gaps ([7], [8]) and other parameters on the perceived personality and/or vocal traits of a speaker, such as dominance or charisma. Despite this interest, to our knowledge, there remains limited research examining the influence of varying levels of suprasegmental amplitude variation on personality trait perception, particularly within the realm of dominance. Here we are focused on amplitude differences, with amplitude and intensity being classified as acoustic parameters while loudness is classified as the corresponding perceptual parameter. We are interested how different linguistic

domains such as *sentence/phrase*, *word* or *syllable* level would influence perception of dominance.

1.1. Previous Research

Previous research has found that differences in (1) amplitude ([9], [10]), (2) loudness ([11]), and (3) intensity ([12]) impact listeners' perceptions of dominance for a given speaker. Results from Tusing & Dillard [13] found increased overall mean vocal amplitude is positively associated with increases in dominance ratings (here defined as a more general concept of dominance) in the production of spontaneous speech. In that study, speech productions were short audio messages which were examined for amplitude differences over several phrases. Due to the short nature of audio length, the experimental design essentially examined amplitude variation perception at sentence level.

Other research has found that more confident individuals speak with greater amplitude ([14]). In this study, participants verbally answered a series of questions, and then scored how confident they were in their own answers. The recorded participant audio was converted to amplitude measurements. The amplitude measurements of these verbal responses were compared to the confidence scores participants gave themselves. Since these verbal answers were short, (one to two words), it could be argued that this study was looking at influences of word stress, thus examining the effect of word stress on *social* dominance (as opposed to overall dominance) perception through confidence scores.

Lastly, high signal intensity is associated with perceptions of dominance ([12], [10]). In the study by Aronovitch [12], participants listened to short audio recordings (around 10 seconds) and were asked to judge several personality traits. One of these scales looked at dominance (ranging from dominant to submissive). The presented recorded audio stimuli were several different speakers spontaneously describing a presented image. Results demonstrated that higher signal intensity increases perceptions of dominance. Like the study by Tusing & Dillard [13], intensity levels were examined for the whole phrase, and again had a more general definition of dominance (overall dominance).

Although these studies often do not differentiate between social or physical dominance and generally emphasize amplitude variations at sentence level, results clearly show there is a relationship between ratings of perceived dominance and acoustic amplitude variation.

For the current study, we opted for two different definitions of dominance. We modelled our definition as proposed by Mueller and Mazur [15, p. 570]: “a dominant person tells other people what to do, is respected, influential, and often a leader, while submissive or subordinate people are not influential or assertive and are usually directed by others.”. This definition was later used in a study by Puts et al. [5] which provided the framework for the current study. More specifically, *physical* dominance is a measure of physical behaviours or traits, such as physical strength, which increase an individual’s status within a given hierarchy, while *social* dominance measures the social component of behaviours or traits resulting in a speaker being viewed as influential and respected as a leader. Essentially, for *physical* dominance, speakers are judged whether they are likely to win in a physical fight. For *social* dominance, it is the likelihood that the speaker is a respected leader.

1.2. Aims & Hypotheses

The present study investigates how suprasegmental amplitude variation on different linguistic/phonetic levels affects listeners *social* and *physical* dominance perception of these same speakers, and additionally which of these dominance types are the most salient in high (positive) versus low (negative) listener ratings. Three different linguistic/phonetic levels were chosen to capture numerous possible amplitude variations seen in acoustic phonetics and linguistics in general: sentence-level (i.e., amplitude manipulation of the entire phrase/paragraph), word-level (i.e., one specific word in focus), and syllable-level (i.e., amplitude differences contrasting stressed compared to unstressed syllables).

We have the following hypotheses: First, we assume that increases in amplitude over the whole sentence will increase ratings of dominance for both *social* and *physical* dominance. Second, increasing amplitude differences at word level (focus) should increase ratings for *social* dominance (see [14]), but not necessarily for *physical* dominance. Lastly, for syllable stress, we expect to see an influence on ratings for *both* types of dominance (similar to our expectations for *sentence* level), however, due to the lack of previous research, it is unclear if listener ratings will be affected more positively or more negatively on the *syllable* level.

2. METHODS AND MATERIALS

2.1. Stimuli

As word material we chose 2 paragraphs (consisting of multiple simple sentences) and 4 isolated sentences. Both the sentences and paragraphs were constructed to have a neutral valence (i.e. the connotations of the stimuli had neither positive nor negative valence to prevent any influence of word material on listeners interpretations). Each paragraph was approximately 12 seconds long.

Six native Canadian English speakers (3 female, 3 male) produced the stimuli. 4 were professional voice actors and 2 were Linguistics graduate students (McMaster University). Due to COVID restrictions, the professional voice actors used their own high-quality microphones and adequate recording environments. During the recording sessions, the speakers were monitored and directed via Zoom by the authors of the current study. The 2 graduate students were recorded on a high-quality microphone (Rode NT1A) via Focusrite Scarlet audio interface inside the soundproof booth of the Phonetics Lab at McMaster University. For all recordings, microphone distance was around 10 cm (microphone being horizontally off-centre from the lips) by 45 degrees. Recordings for each stimulus was repeated 3 times, with the best production selected as the perceptual stimulus for the listeners. Prosodic differences were as tightly controlled as possible across speakers and conditions. After recording, the audio samples were screened ([16]) and carefully checked for achieved accuracy and consistency of produced vocal parameters, the absence of undesired voice quality changes (e.g., sentence-final glottalization) and differences in articulation (e.g., different speech rates and hyper/hypoarticulation). Additionally, a high pass filter (80 Hz for male speakers and 150 Hz for female speakers; 24dB/octave linear-phase) was applied to remove and attenuate any additional low frequency noise which may have been a part of the original recordings. Before amplitude manipulations, the loudness for all stimuli per condition ((i.e., applied to each sentence stimulus or each paragraph) was normalized to 65 dB using the Praat (Boersma & Weenink [17]) intensity normalization procedure.

2.2. Amplitude Manipulations

Each sentence/paragraph’s amplitude was then manipulated on three levels: (1) sentence stress, (2) word stress (*focus*) and (3) syllable stress (i.e. stressed versus unstressed syllable amplitudes). Each of the three manipulation levels was performed on the baseline audio file (i.e. the speaker-produced audio

file for each speaker and each paragraph/sentence) in Adobe Audition ([18]), resulting in an *enhanced condition* (i.e., the *enhanced* condition had increased amplitude values at sentence, word or syllable level). Only for the syllable stress condition we also included a *reduced* condition.

Sentence stress (phrasal stress) was manipulated for the entire sentence or paragraph. The overall sentence intensity was increased by 6 dB. This means that by increasing the intensity on sentence/paragraph level and comparing these manipulations to the originally produced utterance (i.e., baseline) we can examine how the dominance perception of that speaker in the higher intensity condition compares to the perception of the same speaker producing the same sentence/paragraph, but with normal (i.e. -6 dB) intensity as baseline.

Word stress was manipulated by increasing the amplitude of one particular word of each sentence, so that this one word is now in stressed or *focus* position for that sentence, but only with regard to intensity differences (thus not pitch, duration, or vowel quality differences) compared to the rest of the sentence. The selected focus word which was increased in amplitude was selected before the manipulation and was identical for all speakers.

Lastly, *syllable stress* intensity, or more precisely the intensity difference between stressed and unstressed syllables, was manipulated in two ways: an *enhanced* condition (as in the previous two descriptions) and additionally a *reduced* condition, in order to investigate the effect of varying stressed/unstressed syllable intensity differences on listener dominance ratings. For the *enhanced* condition, the intensity of all syllables in stressed positions within a sentence/paragraph was increased by 6 dB, while all unstressed syllables remained unaltered. For the *reduced* condition, both the intensity of stressed syllables was reduced, *and* the intensity of unstressed syllables was increased. We aimed for this split approach (decrease of stressed syllable intensity plus increase of unstressed syllable intensity) to generate the most natural-sounding production for this condition. The *reduced* condition is meant to give the effect of a speaker with diminished distinctions of stressed/unstressed syllable intensity. Stimuli were not repeated, and the final acoustic stimuli count was 180.

2.3. Experimental Setup

The current study explored an academic setting with a focus on dominance ratings for professors by students. The statements used were modelled from the Puts et al. [5] study which asked listeners to judge if a speaker was more likely to win a physical fight

(classified as *physical* dominance) and whether the listener was “extremely dominant” or “extremely submissive” (*social* dominance). In addition to these 2 statements, we created and added two additional scales (one social and one physical). Participants were presented acoustic stimuli through headphones containing either one paragraph or one sentence. Simultaneously, participants were presented with a screen with 2 statements (one for physical and one for social dominance), paired with a sliding scale for each statement. On a following screen, 2 different statements were presented (again, one social and one physical) along with the repetition of the exact same audio stimulus from the previous screen. Participants were asked to listen to the audio stimuli and rate the statements on a continuous sliding scale. In sum, listener ratings were based on a duplet of two statements, relating to whether the speaker was likely to win in a physical fight or whether a speaker was a respected leader. The responses gave a rating on a scale from 0% (strongly disagree) to 100% (strongly agree).

71 participants took part in the perception study. It was conducted online (due to pandemic restrictions) via Prolific, participants reported normal hearing and cognition and were compensated for their time (around 30 minutes). Ethics clearance was obtained from the McMaster Ethics Board (MREB).

3. RESULTS

Since our data was not normally distributed, we conducted multiple Wilcoxon signed-rank pairwise significance tests ([19], [20]) to determine if each manipulated amplitude condition is significantly different from the corresponding unaltered baselines, for both types of dominance. We corrected for multiple comparisons by adjusting p-level thresholds (Bonferroni correction).

We found significant rating differences for the *enhanced* conditions (compared to baseline) at both *sentence* level (physical dominance: $p < .025$ **, social dominance: $p < .025$ **) and at *syllable* level (physical: $p = .008$ *, social: $p < .0025$ **). *Word* level and the *reduced* intensity *syllable* condition were however not significantly different from their baselines. Figure 1 illustrates these findings with a mean plot.

In the figure, it can be seen that for the significantly different *sentence* level condition, participants, as expected, showed higher dominance ratings for the *enhanced* condition compared to baseline. However, the opposite pattern was found for the *syllable* level: Here the *enhanced* condition was rated less dominant (or to turn it around: the baseline condition was rated more dominant).

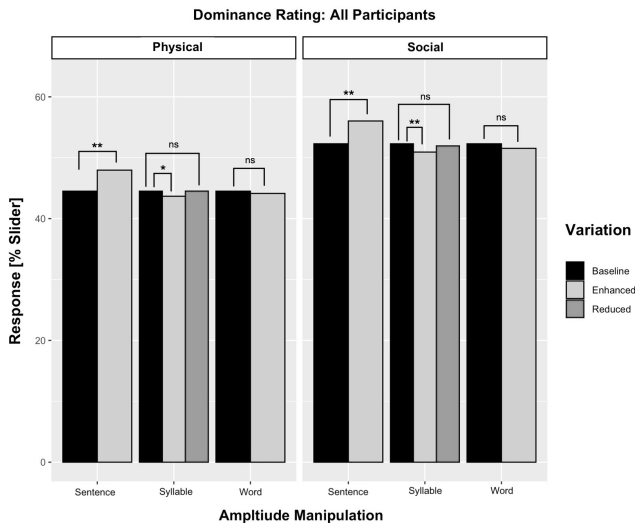


Figure 1: Mean dominance ratings over all speakers and all listeners, split by *physical* versus *social* dominance. The y-axis displays slider position percentages with 0% corresponding to the left extreme value and 100% as the right extreme value. The x-axis shows the different amplitude manipulations levels (*sentence*, *syllable*, *word*)

We found opposite directions for the significantly different dominance ratings of our listeners: Enhancing amplitude differences at *sentence* level significantly increased dominance ratings, whereas enhancing amplitude differences at *syllable* level significantly *decreased* dominance ratings. Manipulations at *word* level or the reduced condition at *syllable* level did not (significantly) affect listeners' dominance ratings.

4. DISCUSSION

Not all amplitude increases lead to higher dominance ratings, the effect strongly depends on the actual linguistic level manipulated. Also, the two examined dominance domains (*physical* versus *social*) show similar patterns across all three linguistic domains.

For the high-level *sentence* condition, our data confirm the results of previous research: increasing the overall amplitude increases the perceptual ratings of dominance [12], [13]. This is true for both dominance domains, and shows that overall increases in sentence level amplitudes positively correlate to increases in perceived *social* and *physical* dominance.

For the mid-level *word* condition, our results contradict results from [14] which found increases in amplitude for *word* stress result in higher scores within the *social* dominance environment. Our data did not confirm these differences for the social dominance domain, and we furthermore also did not find differences for *physical* dominance.

With respect to the two low-level *syllable* conditions, we are not aware of previous literature

examining differences on this level. Our results show, for both the *social* and *physical* dominance domain, that enhancing amplitude differences between all stressed and unstressed *syllables* leads to significant dominance rating differences, but opposite to our findings for the *sentence* level: the *enhanced* condition significantly lowers dominance ratings compared to the baseline, whereas the *reduced* condition had no effect on dominance. It is of interest to note that the effect size for the *syllable* level condition is much smaller than *sentence* level condition (see figure 1), suggesting intensity changes on *sentence/phrase* level are more salient and robust compared to *syllabic* level changes.

5. CONCLUSION

We manipulated amplitudes on three different linguistic levels (sentence, word, syllable) and asked listeners to rate these stimuli using dominance scales with continuous slider scales. We found that our *enhanced* amplitude conditions significantly influenced listener ratings for *sentence* and *syllable* level (but not *word* level), and there was no difference comparing *physical* and *social* dominance domain ratings. However, the direction of dominance ratings was different when comparing *sentence* to *syllable* level. One explanation for the unexpected lower dominance ratings for the *enhanced syllable* condition could be due to increased attention: Unlike sentence or focus conditions, stressed syllable manipulations/variations across an entire utterance are not common. This amplitude manipulation may have marked it more auditorily salient to listeners causing decreased dominances ratings for this level.

Finally, with respect to the more salient and pronounced effect found on *sentence* level, it is worth noting that in real-life conditions listeners will not commonly judge identical speech/word material to compare matched increased versus normal amplitude conditions for full sentences and paragraphs. Thus, possible aims to increase a speaker's dominance by simply increasing overall spoken amplitude levels for a full sentence would not succeed, since the matched amplitude baseline condition would be missing here. Although the results are interesting and appear to demonstrate increased amplitude for whole utterances increases dominance ratings, without the baseline to compare the increased environment, the results remain limited in terms of real world application.

6. ACKNOWLEDGEMENTS

This work was partly funded by the Canadian NSERC grant RGPIN-2018-06518 to the second author.

7. REFERENCES

- [1] Quené, H., Boomsma, G., & van Erning, R. 2016. Attractiveness of male speakers: effects of voice pitch and of speech tempo. In *Proceedings of Speech Prosody*, 8, 1086-1089.
- [2] Berger, S., Niebuhr, O., & Peters, B. 2017. Winning over an audience—A perception-based analysis of prosodic features of charismatic speech. In *Proc. 43rd Annual Conference of the German Acoustical Society, Kiel, Germany*, 1454-1457.
- [3] Yang, Z., Huynh, J., Tabata, R., Cestero, N., Aharoni, T., & Hirschberg, J. 2020. What makes a speaker charismatic? producing and perceiving charismatic speech. In *Proc. 10th International Conference on Speech Prosody*, 2020, 685-689.
- [4] Rafikova, A., Valueva, E., & Panfilova, A. 2022. Voice and Psychological Characteristics: A Contemporary Review. *Psychology. Journal of Higher School of Economics*, 19, 195-215.
- [5] Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. 2007. Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340-344.
- [6] Rosenberg, A., & Hirschberg, J. 2009. Charisma perception from text and speech. *Speech Communication*, 51(7), 640-655.
- [7] Möbius, B. (2003). "Gestalt Psychology meets phonetics—An early experimental study of intrinsic F0 and intensity," in *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, IPA and UAB, 3, 2677–2680.
- [8] Niebuhr, O., & Fischer, K. 2019. Do not Hesitate!—Unless You Do it Shortly or Nasally: How the Phonetics of Filled Pauses Determine Their Subjective Frequency and Perceived Speaker Performance. In *INTERSPEECH*, 544-548.
- [9] Harrigan, J. A., Gramata, J. F., Lucic, K. S., & Margolis, C. 1989. It's how you say it: Physicians' vocal behavior. *Social Science & Medicine*, 28(1), 87-92.
- [10] Scherer, K. R., & London, H., & Wolf, J. J. 1973. The voice of confidence: Paralinguistic cues and audience evaluation. *Journal of Research in Personality*, 7, 31-4.
- [11] Buller, D. B., & Burgoon, J. K. (1986). The effects of vocalics and nonverbal sensitivity on compliance: A replication and extension. *Human Communication Research*, 13, 126-144.
- [12] Aronovitch, C. D. 1976. The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *Journal of Social Psychology*, 99, 207-220.
- [13] Tusing, K. J., & Dillard, J. P. 2000. The sounds of dominance. Vocal precursors of perceived dominance during interpersonal influence. *Human Communication Research*, 26(1), 148-171.
- [14] Kimble, C. E., & Seidel, S. D. 1991. Vocal signs of confidence. *Journal of Nonverbal Behavior*, 15(2), 99-105.
- [15] Mueller, U., & Mazur, A. 1997. Facial dominance in Homo sapiens as honest signaling of male quality. *Behavioral Ecology*, 8, 569–579.
- [16] Hairer, M. 2021 [software]. Amadeus Pro (version 2.8.8). London, United Kingdom; HairerSoft.
- [17] Boersma, P., & Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. Version 6.2.23, retrieved 8 October 2022 from <http://www.praat.org/>
- [18] Adobe Inc. 2021. *Adobe Audition*. Retrieved from <https://adobe.com/products/audition>
- [19] R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [20] RStudio Team. 2009-2018. *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.