Control experiments on the frequency scale of speech intonation
Hermes, D.J.; Rump, H.H.

Published in:
Journal of the Acoustical Society of America

Published: 01/01/1994

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 02. Jan. 2019
The fundamental frequency \( (F_0) \) of speech has been used for examining both transient emotional states [D. A. Vakoch, J. Acoust. Soc. Am. 95, 2974 (1994)] and more stable personality traits. For example, low perceived pitch has been associated with attributes of high emotional stability [K. R. Scherer, Eur. J. Soc. Psychol. 8, 467–487 (1978)]. In the present study, characteristics derived from \( F_0 \) were related to five personality traits as measured by the NEO-PI: neuroticism, extraversion, openness, agreeableness, and conscientiousness. A high degree of openness was predicted by a high mean \( F_0 \) of an utterance \((p<0.025)\) and by a great variability in \( F_0 \) within an utterance \((p<0.004)\). The best predictor of neuroticism was peak \( F_0 \) within an utterance, with subjects scoring high on this scale having a higher maximum \( F_0 \) \((p<0.01)\). The latter finding, in conjunction with previous research on attributes of emotional stability, suggests a convergence between objective acoustical measures and subjective perceptions of personality.

The relative timing of consonants and vowels of a syllable can be affected by factors such as speaking rate, focal stress, and distinctive vowel length. Speaking rate and focal stress tend to have a global effect on syllable-internal timing, whereas the effect of distinctive vowel length in Norwegian is generally believed to be localized to the syllable rhyme. When these factors converge on a syllable, the syllable-internal timing simultaneously reflects their concurrent effects. The goal of the present study is to examine and characterize the concurrent effects of speaking rate, focal stress, and distinctive vowel length on syllable-internal timing of C1VC2 components. Conversations were developed in which 12 Norwegian CVCs containing /i, o, a, i, o, a/ were either focused or nonfocused by the discourse. The set of conversations was produced at three speaking rates by native speakers of Norwegian. Results indicate that speaking rate and focal stress have an overall affect on the timing of components in the syllable. In addition, the effects of distinctive vowel length and focal stress on components of the syllable rhyme were found to be interdependent across speaking rates. These findings are discussed in terms of syllable- and rhyme-internal timing.

When bearing a word which is continuously repeated, listeners report that the word seems to change into different forms and then vacillates among these forms. This auditory illusion has been called the verbal transformation effect [K. M. Warren, Br. J. Psychol. 17, 249–258 (1961)] and is considered to be caused by the lack of verbal context, produced by the repetition. This study further investigated the role of context in the interpretation of auditory stimulation. Twenty-six listeners heard the individual, continuous, repetition of six stimuli: an English sentence, an English word, and a Chinese word (each played forward and backward). It was assumed that listeners will report verbal transformations with all stimuli except the English sentence, since only this condition contained sufficient contextual cues. Results showed that both forms of the English sentence were treated in the same manner, and that significantly more verbal transformations were reported with the words.

According to Hermes and Van Gestel [J. Acoust. Soc. Am. 90, 97–102 (1991)], prominences of accent-lending pitch movements in different registers are equal when their excursion sizes are equal on the ERB-scale rate. A reanalysis of their experiments showed that an octave relationship between the high pitch levels in the different registers may have been crucial. In order to test whether this factor influenced the outcome of their experiments, adjustment experiments were performed with stimuli in different low and high registers. The results, supporting the ERB-rate scale hypothesis of Hermes and Van Gestel, showed that the octave relationship had no systematic influence on the results. In an additional experiment, stimuli were resynthesized from a male and a female voice in order to test the effect of the speaker’s natural pitch range. The male and female stimuli gave about the same results. In sum, the effect of pitch excursion on prominence is very robust. An unexpected finding was that, on average, the excitations of the pitch movements in the high register were about 0.15 E (0.5–1 semitone) larger than those of the pitch movements in the low register when lending equal prominence.

According to the model proposed by Hermes and Rump [J. Acoust. Soc. Am. (in press)] the prominence lent by accent-lending pitch movements is proportional to the pitch-level difference between the vowel nucleus of the accented syllable and the vowel nucleus of the preceding syllable. The pitch level on the lower declination line or baseline and the pitch level on the upper declination line or topline are thus regarded as equally important. By varying the height of the pitch levels on the baseline this assumption was tested against the alternative assumption that listeners’ perceptual evaluation of each of the stimuli when adjusting prominences to be equal. Results showed that pitch levels on the toplines were adjusted to be more or less equal, but that the pitch-level difference between the low and the high pitch level influenced the adjusted height of the high pitch level, too. It is concluded that the proposed model makes claims about the relevance of the position of the low pitch level which is too strong. The role of pitch levels in relation to the perception of prominence will be discussed here.