"Minor Third, Who?": The Intonation of the Knock-Knock Joke

Jeremy Day-O'Connell

Department of Music, Knox College, Galesburg, IL

jdayocon@knox.edu

Abstract

In an effort to examine the intonational phenomenon of stylized intonation, knock-knock jokes were collected and phonetically analyzed. Results showed intonation that varied considerably from subject to subject but which was nevertheless constrained in a way that supports a hitherto unexamined supposition among musicologists and linguists: that stylized intonation is defined by the musical interval of a minor third. Results showed a preference for intervals approximating a minor third, as well as an unexpected "boundary" role for the minor third itself, which is interpreted as a consequence of physiology.

Index terms: intonation; stylized fall; calling contour; music and language; minor third; jokes; knock-knock joke

1. Introduction

Music and language are universal activities of human expression that share a host of basic sonic resources (including pitch, rhythm, and timbre) and perceptual and cognitive faculties (including pattern matching, categorization, syntax and semantics). These and other similarities have often led to the speculation that music and language were closely related in the evolutionary history of our species--two branches of a single, primal form of human communication. Scientists have just begun to study the precise relationship between music and language; indeed, recent work has spawned a veritable discipline unto itself, unifying subfields from musicology, linguistics, and psychology. (See, for instance, [1].)

Despite the commonalities mentioned above, and notwithstanding our casual reference to the "melodiousness" of a speaker's voice or the "eloquence" of a musical composition, it would seem self-evident that music and language are (at least these days) ontologically separate. It is therefore noteworthy that speakers of English occasionally adopt a particular intonational mannerism that resembles the use of *sung*, as opposed to spoken, syllables: playful or attention-seeking speech intoned in deliberate, discrete dyads, as in "Yoo-hoo" or "Bye-bye." This contour is chiefly used in infant-directed speech but is regularly recruited in adult-directed speech for endearing (or mock-endearing) effect; it is also used in calling at a distance ("Din-ner!"), taunts ("Nya-nya!"), and in *ad hoc* group chants (as in the American basketball jeer, "Air-ball!").

Linguists have described this phenomenon variously, as "stylized intonation" [2], the "calling contour" [3, 4], "vocative chanting" [5], and "stereotyped language" [6] and have, moreover, noted its use in several linguistic cultures [7]. It has also been cited (albeit only casually) by many musicologists, who concur with vague claims of universality [8–12]. Importantly, observers (whether linguist or musicologist) inevitably cite a specific musical interval in connection with this linguistic formula: the minor third.

Despite the widespread acknowledgement of stylized intonation and the provocative suggestion of a musical universal (the minor third), however, scholars have never studied this phenomenon systematically but have contented themselves with largely anecdotal evidence. The lack of serious literature on the topic can perhaps be explained by its precarious location at the intersection of disparate disciplines. The lack is deeply regrettable, however, as careful investigation would likely bear important implications for more fundamental questions in anthropology and evolutionary psychology, such as the apparent existence of similarities among the musical scales of the world (albeit very rough similarities [13]), and indeed, speculations about the origins of human language and music-making.

The present study then represents the first empirically based description of a neglected phenomenon located squarely, and fortuitously at the intersection of music and language: the "stylized fall," described by Ladd as "a 'stepping down' from one sustained tone to another" [2]. The intention is to instrumentally measure the English stylized fall, as a first step toward assessing the widespread supposition that the minor third bears a special, and universal, linguistic function.

2. Method

The study of prosody is typically hindered by a set of methodological problems, chiefly the difficulty of collecting "ecologically valid" utterances. Stylized intonation is perhaps especially susceptible to this difficulty, as it is in a way inherently "artificial," tantamount to a sort of performance. The present study overcomes this difficulty by targeting a particularly spontaneous mode of speech: joke-telling. The "knock-knock" portion of knock-knock jokes was taken to be a reasonable example of the stylized fall, resembling as it does (in both form and function) the "calling" category of stylized intonation ("Yoo-hoo," "Johnny," etc.).

Sixty-three college students (ages 17 to 23) were recorded telling a knock-knock joke (from their own recall). Two conditions were utilized, in order to assess the effect of the "calling" mode of delivery often associated with the stylized fall: 22 of the 63 subjects told their jokes to the experimenter across a long (8-foot) table (the "distance" condition); the rest of the subjects were seated immediately beside the experimenter (the "normal" condition).

Using phonetic analysis software (Praat), each "knock-knock" was isolated, segmented into two syllables, and analyzed for fundamental frequency (F0). Of the 63 collected utterances, 6 were found to use a rising contour, and 6 yielded insufficient pitch tracks due to background noise and idiosyncrasies of the subject's delivery. Fifty-one utterances were therefore used in

the final analysis: 33 in the normal condition (12 male, 21 female) and 18 in the distance condition (10 male, 8 female).

This analysis consisted of calculating the following within-subject data (see Figure 1): *mean pitch* for each of the two syllables (high "knock" and low "knock") using the geometric mean of F0 (equivalent to mean semitone-scaled pitch); *pitch stability* for each syllable using the standard deviation; and *interval size* for the utterance using the ratio of mean high pitch to mean low pitch (converted into semitones). Pitch stability and interval size then formed the basis for between-subject generalizations.

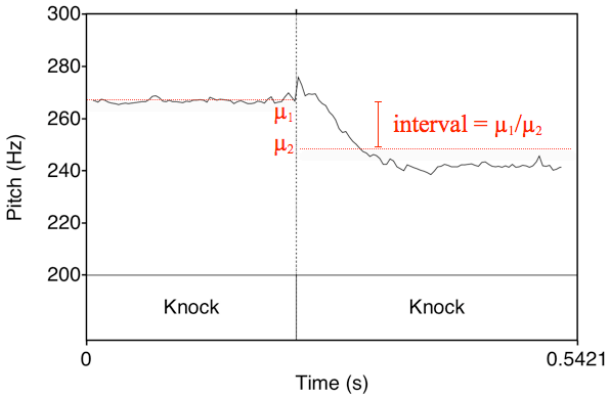

Figure 1. The pitch plot of a typical utterance segmented into its two syllables, with schematic illustration of analysis.

3. Results

The mean interval size was found to be 2.16 semitones (standard deviation=1.46 st), which corresponds to a musical interval slightly larger than a major second. A distance effect was also observed: the mean interval size in the distance condition was found to be 2.77 semitones (standard deviation=1.90 st), which corresponds to a musical interval slightly smaller than a minor third; this was 65% larger than the normal condition's mean interval size, 1.80 semitones (standard deviation=1.03 st), which corresponds to a musical interval slightly smaller than a major second (Fig. 2).

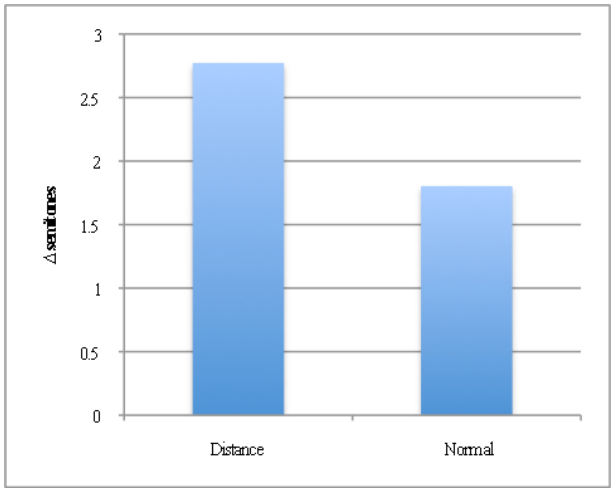

Figure 2. Mean interval size of "knock-knock"

Pitch stability also showed an effect of distance, though chiefly for the lower of the two "knock-knock" notes: the mean standard deviation for the high note was 0.58 st in the normal condition and 0.54 st for the distance condition; for the low note it was 0.73 st and 0.30 st, respectively (Fig. 3).

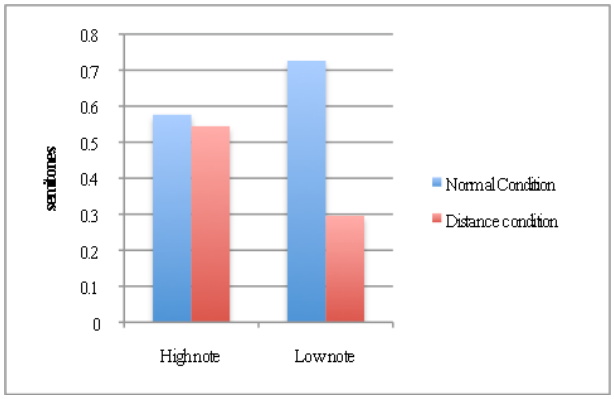

Figure 3. Mean standard deviation of "knock-knock" notes

Most interestingly, the interval of a minor third (3 semitones) represented an upper limit on interval size for most utterances: 46 of the 51 utterances (90.2%) used interval sizes up to a minor third. And as indicated in the histogram in Fig. 4, the most common interval sizes were those just smaller than a minor third.

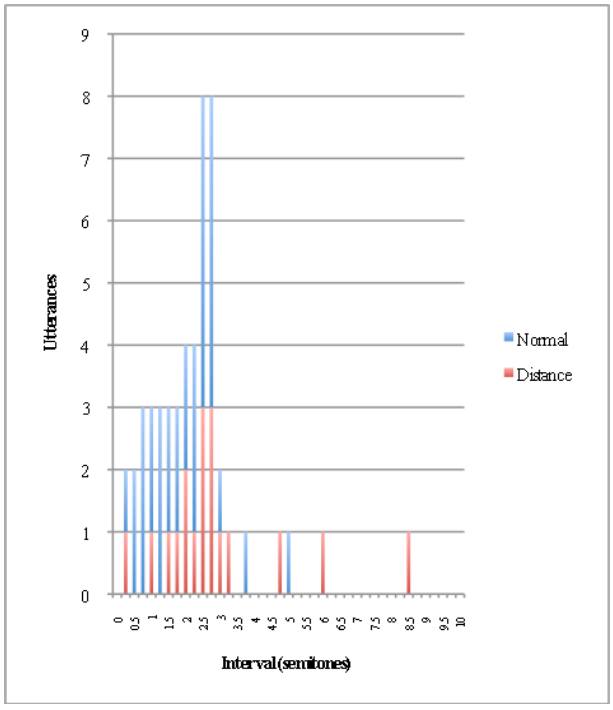

Figure 4. Histogram of knock-knock intervals

4. Conclusions

The conclusions of this study are limited by the circumscribed and idiosyncratic nature of the sampled speech (the "knock-knock"). Nevertheless, the phonetic regularities that emerge

may be typical of stylized intonation in English, something further study could confirm. Most importantly, notwithstanding the undeniable variety of intonational realization observed, an interval of between 2 and 3 semitones does appear to be preferred by speakers, and the minor third itself appears to constitute some sort of boundary for most speakers. This boundary could be understood as physiologically derived: perhaps intervals beyond the minor third are categorically less comfortable--more effortful--to produce. (The observed distance effect could be similarly explained with respect to vocal effort.) This possibility, which could be tested experimentally, gives promise to the hypothesis (only implicitly expressed by some of the writers referenced above [7–12]) that the minor third transcends language-specific phonology and constitutes an intonational phenomenon of cross-cultural importance.

Future studies could apply the present method to other English-speaking communities as well as to other language communities that have knock-knock jokes (e.g., France, South Africa, Morocco). Other studies of stylized intonational phenomena throughout the world could also be compared.

5. Acknowledgements

This research was supported by the Dean's Office of Knox College.

6. References

[1] Patel, A. (2008). *Music, Language, and the Brain*. New York: Oxford.

[2] Ladd, D. R. (1978). "Stylized Intonation." *Language* 54/3, 517-40.

[3] Ladd, D. R. (1996). *Intonational Phonology*. New York: Cambridge.

[4] Abe, I. (1962). "Call-Contours." In Antti Sovijärvi, ed., *Proceedings of the Fourth International Congress of Phonetic Sciences*. The Hague: Mouton, 519-23.

[5] Liberman, M (1979). *The Intonational System of English*. New York: Garland.

[6] Chao, Y. R. (1956). "Tone, Intonation, Singsong, Chanting, Recitative, Tonal Composition, and Atonal Composition in Chinese." In Morris Halle, ed., *For Roman Jakobson. Essays on the Occasion of His Sixtieth Birthday*. The Hague: Mouton, 52-9.

[7] Various authors in Hirst, D. and Di Cristo. A., eds (1998). *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge.

[8] Ringer, A. (2001). "Melody." In *The New Grove Dictionary of Music and Musicians*, 2nd ed., edited by Stanley Sadie and John Tyrrell. London: Macmillan, 16: 363-73.

[9] Alper, C. D. (1992). "Early Childhood Music Education." In *The Early Childhood Curriculum: A Review of Current Research*, edited by Carol Seefeldt. New York: Teachers College Press, 237-63.

[10] Heaton, C. (1992). "Air Ball: Spontaneous Large-Group Precision Chanting." *Popular Music and Society* 16, no. 1: 81–83.

[11] Brailoiu, C. (1984). *Problems of Ethnomusicology*, trans. and ed. A. L. Lloyd. New York: Cambridge.

[12] Bernstein, L. (1976). *The Unanswered Question*. Cambridge, MA: Harvard University Press.

[13] Nettl, B. (2000). "An Ethnomusicologist Contemplates Universals in Musical Sound and Musical Culture." In Nils L. Wallin, Björn Merker, and Steven Brown, eds., *The Origins of Music*. Cambridge, MA: MIT Press, 463-72.