Speech Perceptual Periodicities Facilitate Lexical Access and Contextual Emotional Responses

Alfonso Ceccherini-Nelli[1], Marta Kulinska[2], David Murray[3] & Harry M. Logan[4]

[1] Alberta Hospital Edmonton. Addiction and Mental Health, Alberta Health Services, Canada

[2] Stratford School, Cogito Program, Canada

[3] Phi Mu Alpha Sinfonia, USA

[4] English Language and Literature, University of Waterloo, Canada

Correspondence: Alfonso Ceccherini-Nelli, Alberta Hospital Edmonton, Addiction and Mental Health, Alberta Health Services, 17480 Fort Road, Edmonton, AB, Canada. E-mail: alfcn@live.com

Abstract

Here we show that the frequency of consonant sound repetitions in different samples of language describe a continuum, with speeches of outstanding politicians lying on the upper end and writings of schizophrenic subjects on the lower end. Furthermore, we show that the perceptual periodicities of a speech of Obama facilitate lexical access and induce emotional responses analogous to music and congruent with the context. This provides indirect evidence of emotional entrainment between the speaker and his audience. Our findings also support the transformative technology of the mind theory of music (Patel, 2010). In fact, we argue that music and poetry were invented by transforming the previously acquired ability to analyze the prosodic component of language, which was advantageous in social interactions.

Keywords: rhetoric, prosody, entrainment, Obama, schizophrenia, music, poetry, Lexical Decision Task

1. Introduction

Communication signals in primates have evolved to share information (Zuberbuhler, 2012), to communicate the emotional state of the caller (Darwin, 1872), to change the emotional state of the receiver (emotional contagion) (Owings & Morgan, 1998) and ultimately to manipulate the behavior of listeners (Dawkins & Krebs, 1978). In humans, language is essential in all social interactions, particularly those leading to leadership selection. Although the art of effective or persuasive communication (rhetoric) has been investigated since the ancient Greek civilization, we still understand poorly its psychophysiology. Night (2009) has suggested that entrainment-inducing stimuli are important in successful oratory. Entrainment refers to the phenomenon of a rhythm (or something which varies rhythmically) causing another gradually to fall into synchronism with it (Oxford Dictionaries, 2013). Entrainment plays an attention-directing role in music perception and it is believed to have a role in social bond formation as well (Night, 2009). For some (Phillips-Silver, Aktipis, & Bryant, 2010), social entrainment is a category of entrainment characterized by responsiveness to rhythmic information generated by others. In particular, acoustic-prosodic entrainment is predictive of dialogue success between pairs of individuals (Levitan et al., 2012). The occurrence of perceptual regularities has been empirically observed in acclaimed English poetry (Logan, 1976, 1985). The close relationship between political oratory and the rhythms of music were noted by Darwin in "The Descent of Man" (Darwin, 1872). However, robust empirical evidence of the phenomenon is lacking. It also remains unexplained why humans would care for these perceptual repetitions while choosing their leaders. Finally, the relationship between oratory, poetry and music is poorly understood. The primary objective of this study was to analyze empirically the presence of rhythmical repetition of any consonant sounds in renowned political speeches of Obama, Churchill and King Henry VIII and compare it with other writings from local leaders, technical manuals and subjects with schizophrenia—a mental illness characterized by poor social functioning (Pinkham, Penn, Perkins, & Lieberman, 2003) language disturbances (Ceccherini-Nelli & Crow, 2003) inappropriate affect (Moskowitz & Heim, 2011), impaired recognition and expression of emotional prosody (Hoekert, Kahn, Pijnenborg, & Aleman, 2007), and impaired musical perception (Wu et al., 2013; Hatada et al., 2014). We chose two subjects with a known diagnosis of schizophrenia.

The first was the mathematician and Nobel laureate John Nash. We analyzed three of his writings published either before or after the onset of the illness in 1959. The second subject is also high functioning, with only one brief hospitalization for schizophrenia, in spite of never being treated for any significant length of time. We analyzed one of her published books, that we identify here as a "Comic Novel". Secondary objectives were to test if the sound repetition patterns observed in political speeches exerted a positive attention-direction effect, measured with a lexical decision task, and if the underlying rhythm of speech, once isolated from its semantic component and converted to musical notation, could induce contextual emotional reactions in the listeners.

2. Linguistic Analysis

2.1 Methods

Both qualitative and quantitative approaches are available to analyze the patterns of rhythm and acoustic effects used in spoken and written language. The qualitative approach consists in outlining the repetition of specific patterns like alliteration, onomatopoeia, consonance, assonance, etc. The quantitative approach measures the frequency of certain sound patterns considered specific to poetry.

Logan (1976, 1985) devised a methodology to measure quantitatively sound effects in the English poetry. In a nutshell, Logan (1985) proposed to analyze the frequency and pattern of repetition of individual phonemes to form themes and motifs.

In the first experiment, selected texts were separated into lines and then phonetically transcribed into the IPA alphabet. We studied sound patterns and motifs by analyzing the frequency of different consonant sounds in the entire text, and the repetitions of consonant sounds across lines. We compared these frequencies between different writing samples using chi-square and t statistical tests. In the first experiment, texts were separated into lines according to either the long pauses made by the speaker, whenever an audio recording was available, or punctuation. This allowed us to differentiate the clustering of sounds together in proximity from the total number of those sounds whenever they occur. The phonetic transcription was made using the software PhoTransEdit (Version 1.7d. downloaded from www.photransedit.com). Words left without a transcription were converted to the IPA alphabet using the Oxford Advanced Learner's Dictionary. A manual check of the presence of homographs was also performed.

The choice of using a computer automated phonetic transcription was dictated by practicality. It is commonly accepted that the phonetic transcription provided by Oxford Dictionaries is the most reliable. However, the manual input of data and the manual counting of the frequency of individual IPA symbols would have required time and efforts beyond the scope of this paper. Nevertheless, we tested the level of concordance between the phonetic transcriptions obtained with PhoTransEdit and the Oxford Advanced Learner's Dictionary randomly selecting a paragraph of the speech of Obama. The concordance using these two methods was very high (Cohen's Kappa=0.95; 95% C.I. 0.93 to 0.98). Therefore, we believe that the results reported in this paper would not be significantly different if the transcription had been made manually using the Oxford Dictionaries.

2.2 Results

Results are summarized in Tables 1 and 2. The frequency of dental sounds was the highest in Obama, the Medical Director, and Nash (Nash, 1958)—(71% in all three) and significantly higher than the technical manual (68%, χ^2=5.80, P=0.016) and the Comic Novel (68%, χ^2=8.70, P=0.0032). The number of repetitions of any consonants across lines (mean±SD) was the highest in Obama (16.2±5.1) and the lowest in the Comic Novel (8.8±6.0; t(d)=-7.36, P<0.0001). The number of repetitions of consonants for Nash was 13.5±4.2 in his 1950 paper, 15.2±5.5 in his 1958 paper (NS in the comparison with Obama) and down to 10.7±4.2 in his 1966 paper (Nash, 1966) in the comparison with his 1950 paper (Nash, 1950): t(d)=3.4, P=0.0013. After adjusting for the length of the lines, the number of consonant repetitions was still significantly higher in Obama than Nash (Nash, 1966) (t(d)=-0.58, P=0.02) and the Comic Novel (t(d)=-1.21, P<0.0001).

The analysis of distribution of the 5 dominant sounds in each text repeated across lines confirmed that Obama had the highest number of repetitions, followed by Nash (Nash, 1950 & 1958), i.e., before the onset of schizophrenia and the Comic Novel had the lowest (χ2=102.43; P=0.000002).

Table 1. Political speeches and other prose—summary of findings

	Dental Sounds %; difference with Obama: χ^2, p	Any consonant repetitions across lines. Mean (SD). Difference with Obama: t-test, P	Any Consonant repetitions across lines. Mean (SD) adjusted for line length. Difference with Obama: t-test, P
B.Obama—Prague speech (2009)	71%; NA	16.2 (5.1) NA	7.7 (2.0) NA
W.Churchill—We shall fight on the beaches speech. (1940)	69%; χ^2=7.52, P=0.0061	11.7 (4.2) -11.665, P<0.0001	7.1 (2.1) -3.399, P=0007
Henry VIII—Speech to Parliament (1545)	69%; χ^2=3.23, P=0.0721	12.2 (4.6) -7.498, P<0.0001	7.1 (2.1) -2.67, P=0079
CEO Message	70%; χ^2=0.446, P=0.504	10.6 (3.3) -8.805, P<0001	6.8 (2.2) -3.151, P=0.0018
Medical Director Newsletter	71%; χ^2=0.142, P=0.706	13.7 (1.8) -3.18, P=0.0016	7.3 (3.4) -1.334, P=0.1832
Technical Manual	68%; χ^2=5.80, P=0.016	12.7 (7.34) -3.91, P=0002	7.6 (3.1) -0.29, P=0.7725
Nash (1950)[1]	68%; χ^2=2.59, P=0.1078	13.5 (4.2) -3.043, P=0.0026	7.7 (2.0) 0.172, P=0.8639
Nash (1958)[1]	71%; χ^2=0.073, P=0.786	15.2 (5.5) -1.443, P=0.1499	7.5 (2.1) -0.728, P=0.4671
Nash (1966)[2]	70% χ^2=1.19, P=0.276	10.7 (4.2) -5.5147, P < 0.0001	7.1 (2.2) -0.5809, P=0.0222
Comic Novel (2011)[3]	68%; χ^2=8.702, P=0.0032)	8.8 (6.0) -7.3648, P<0.0001	6.4 (2.8) -1.2180, P<0.0001

[1] Written before the onset of schizophrenia.

[2] Written after the onset of schizophrenia.

[3] Published by a subject previously hospitalized with a diagnosis of schizophrenia.

Table 2. Distribution of sound patterns repeated across lines

	5's	4's	3's	2's	1's	0's
Obama Prague (2009)	68%	26%	5%	1%	0%	0%
Nash (1950) [1]	68%	16%	16%	0%	0%	0%
Nash (1958) [1]	57%	38%	5%	0%	0%	0%
Henry VIII (1545)	54%	34%	9%	3%	0%	0%
Churchill Beaches (1940)	49%	34%	13%	3%	0%	0%
CEO message	48%	33%	16%	2%	0%	0%
Technical Manual	58%	23%	12%	4%	4%	0%
Nash (1966) [2]	44%	37%	15%	4%	0%	0%
Med. Dir. Newsletter	49%	30%	14%	4%	5%	0%
Comic Novel (2011) [3]	41%	29%	18%	9%	3%	1%

Percentage of lines with repetition of the 5 dominant consonant sounds (5's meaning that all 5 dominant consonant sounds are found in a line, 4's stands for 4 of the 5 dominant consonant sounds are found in a line, etc.). The different prose samples are ranked from the highest frequency of sound repetitions (Obama) to the lowest one (Comic Novel) using a composite score.

Chi-square=102.43; d.f.: 45; P-value=0.000002

5s versus Less than 5s: Chi-square=31.282; d.f.: 9; P-value=0.000265

[1] Written before the onset of schizophrenia.

[2] Written after the onset of schizophrenia.

[3] Published by a subject previously hospitalized with a diagnosis of schizophrenia.

3. Lexical Decision Task

3.1 Methods

In the second experiment we used a lexical decision task (LDT) paradigm to test if the sound repetitions found in Obama's speech would result in better lexical access than the Comic Novel, presumably due to a phonological priming effect.

LDT is an experimental task in which subjects have to decide as fast as possible whether a given string of letters is a word or a non-word. Reaction times and accuracy of responses (% of correct answers) are the two variables measured in LDT. Reaction times and accuracy are improved by preceding the target word with some specific words (primers) which relate semantically (semantic priming) or phonologically (phonological priming) (Rouibah, Tiberghien, & Lupker, 1999; Apfelbaum, Blumstein, & McMurray, 2011).

The lexical-decision task computer-aided tests used in this experiment were developed using Inquisit 3 package (2012). In each test, 100 consecutive words and non words were presented on the screen of a laptop. All words (nouns, adjectives, verbs, adverbs) were consecutively taken from either the Obama's Prague speech or the "Comic Novel" written by a person with a history of schizophrenia. 25% of these words had been modified by adding a syllable in the middle of the word (non-words). We administered two different lexical decision task computer aided tests to 18 volunteers, 8 males and 10 females, 40.3±16.2 years (mean±SD) of age, using words from Obama in one test and words from the Comic Novel in the other one. Volunteers were instructed to press different keys on a keyboard after the display of either a word or non-word. Half the subjects were randomly selected to first complete the lexical decision task with Obama's words, followed by the task with the Comic Novel words. The other half of volunteers completed the two tasks following the opposite order. This was aimed to reduce any bias caused by training.

3.2 Results

The mean±SD accuracy in correctly identifying word and non-words was 94%±24% with words taken from Obama's speech and 92%±28% with words taken from the Comic Novel (Fisher's Exact Test; P=0.0209). The mean±SD reaction times in providing a response was 436±339 msec. using words from Obama's speech and 488±1046 msec. using Comic Novel words (Welch test for unequal variances: 1.974, P=0.0485).

4. Emotional Perception of Music

4.1 Methods

In the final experiment, we converted the prosodic information in a section of Obama's speech and of the Comic Novel to musical notation (see fig.1 for an example) and asked 22 volunteers (6 males and 16 females, mean±SD age 44.8±12.1 years) to rate the tunes, played with a synthesizer, using a modified version of the Geneva Emotional Music Scale (GEM-9) (Zentner & Eerola, 2009). Volunteers were asked to rate two other songs. One song ("Annoying") was a compilation of tunes played with a flute, downloaded from You Tube, using as search keywords "annoying", "disturbing" and other negative terms. The last song played to volunteers was the Bird Catcher's aria from Mozart's Magic Flute opera. The volunteers were unaware of how each song had been composed, although several volunteers inevitably recognized Mozart's music. In this experiment, the hypothesis tested was that the music derived from Obama's speech would evoke emotional responses congruent with the semantic content. For the music evoked by the Comic Novel we expected emotional responses less congruent with the declared comic/joyful intent. The "annoying" and Mozart's music pieces were added to the comparison to improve assay sensitivity and we hypothesized that they would evoke negative and positive feelings respectively.

Figure 1. Excerpts from Obama's Prague speech, with text and IPA phonetic and musical notation transcriptions

The Geneva Emotional Music Scale (GEM-9) (Zentner & Eerola, 2009) was modified by adding a tenth item, which we called "Chaos" defined as dysfunctional, troubled, choppy, awkward, and indecisive. Each emotional experience of the GEM was rated from 1 (not at all) to 5 (very much).

The vocal tract is an open cavity including the nasal cavity, mouth cavity, pharynx and larynx. It ends at the vocal folds, which are made of soft skin-like tissue. The voice is like other wind instruments. There is a sound generating device and a resonant cavity (Parker, 2001).

Continuous speech is a set of complex audio signals which consist of a fundamental frequency (F0) and its harmonic components (formants F1, F2, F3…) produced by the vocal folds (Lemmeti, 1999).

Several methods have been proposed in the literature to analyze the prosodic information of speech. The most commonly used method is to automatically display the F0 contour of speech samples with computer programs like Praat (www.Praat.org) or WinPitch Pro (http://www.winpitch.com). However, also the F1 and F2 contours are useful in discriminating different emotions expressed in speech (Yildirim et al., 2004) and in language recognition systems (Martinez, Lleida, Ortega, & Miguel, 2013).

We are aware of two previous methods for turning text into music. The P22 Music Text Composition Generator (Kahn, 2007) allows any text to be converted into a musical composition. In this program every letter of the alphabet is assigned to a note on a scale. This allows for any text to be converted into musical notation. However, no consideration is given to the pronunciation of letters in the context of different words. Furthermore, the exact rules of converting an alphabet letter to a particular note are not publicly available.

The Listening Machine method (Jones & Gregson, 2012) achieves this objective by mapping syllables to pitches of a scale. Consonants are dropped, leaving only the vowel sounds to be transformed to musical notes based on their typical fundamental frequency for an English speaker. The Listening Machine method preserves the rhythms and dynamics introduced by punctuation and stress. We could not consider using this method because our empirical findings about perceptual repetitions refer exclusively to consonant sounds, which are dropped in the Listening Machine.

In this study we propose a new method which takes elements from the two projects, mentioned above. Our method of conversion of text to musical notation firstly involved the phonetic analysis of Obama's voice spectrogram, using the program Praat (version 5.3.10; retrieved from www.praat.org). For each IPA symbol we measured the average frequency of the fundamental and first 3 formants. By comparing the different frequencies, we observed that the first formant of Obama's voice had the widest range of frequencies across the IPA, compared to the fundamental and other formants. Therefore, using the frequency of the 1st formant (F1) only, one of the authors (AC_N) converted each IPA symbol of the phonetic transcriptions to the closest musical note (Suites, 1998). For example, "S" was converted to C6 musical note, because the F1 of "S" was on average 1088 Hz. "P" was converted to F5 note, because the F1 of "P" was on average 697 Hz. Using this method the original phonetic transcriptions of a section of Obama's speech and of the Comic Novel were transformed into a series of musical notes. Finally, these series of notes were converted to musical notation by one of the authors (DM), who remained unaware of the authorship of the two pieces. The musical scores were then played with a synthesizer using the sound of a flute. Although untested, this method was the best one we could think of in order to preserve the emotional prosodic information of speech.

4.2 Results

The music created using the prosodic information in the speech of Obama was given by volunteers an average score>2 ("somewhat") on the following GEM items (mean±SD): Nostalgia (2.68±1.43) and Peacefulness (2.36±1.29). For the Comic Novel, Wonder obtained the highest mean score (2.00±1.23) and Joyful activation score was 1.54±0.91. The annoying music was scored high on Chaos (4.18±0.90) and Tension (3.59±1.37). Mozart's music was given high scores on Joyful activation (4.00±1.34), Power (2.68±1.49), Tenderness (2.54±1.40), Peacefulness (2.54±1.40), and Wonder (2.23±1.31). The maximum total score of GEM was observed for Mozart (21.3±7.2) and the lowest for Comic Novel (16.9±5.5), however the differences across groups were not significant (Kruskal-Wallis test=5.58, NS). Differences across groups were significant (Kruskal-Wallis test: P<0.005) for the following items of GEM: Power, Tenderness, Nostalgia, Peacefulness, Joyful Activation, Sadness, Tension, and Chaos (Fig. 2). As shown, for Obama's music, the highest score was on the item Nostalgia, the score on Peacefulness was almost at par with Mozart, and the score on Tension was the lowest across groups. These emotional responses were congruent with the semantic context of Obama's speech on nuclear disarmament. Comic Novel obtained the highest score on Sadness across groups (Kruskal-Wallis test, P=0.0046), incongruent with the semantic context.
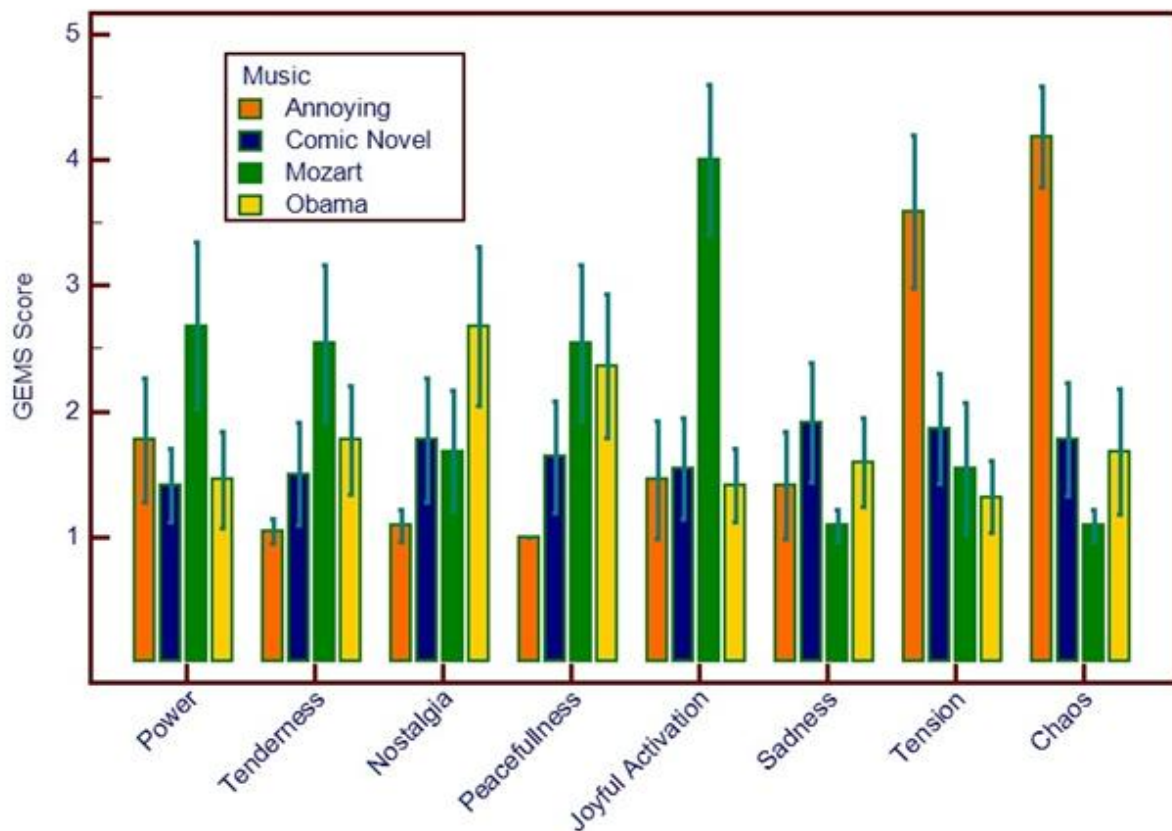
Subjects were asked to rate the intensity of their feelings while listening different samples of music according to this scale: 1-Not at All, 2-Somewhat, 3-Moderately, 4-Quite a lot, 5-Very much.

Figure 2. Items of Geneva Emotional Music Scale (GEMS) (mean±95% C.I.) with significant differences across different music groups (Kruskal-Wallis test: P<0.005)

5. Conclusions

In the first part of this paper, we showed that the frequency of consonant sound repetitions in different samples of language describe a continuum, with speeches of outstanding politicians lying on the upper end and writings of schizophrenic subjects on the lower end of the continuum.

The occurrence of acoustic perceptual periodicities appears to be specific to political speeches. In a parallel study (Ceccherini-Nelli et al., in preparation), contrary to our initial hypothesis that the same perceptual periodicities would be found in any persuasive communication, we failed to find a higher frequency of perceptual periodicities in love letters than in emotionally neutral prose and even break-up letters.

Our findings provide robust empirical evidence that acclaimed political speeches have superior organized rhythmic and melodic patterns compared to other types of verbal or written communication.

In the remaining parts of this paper, we address possible explanations for why such superior organized rhythmic and melodic patterns would make a political speech more persuasive.

Using a lexical decision task (LDT) paradigm, we demonstrated that using the sequence of words in Obama's speech one can achieve shorter reaction times and improved accuracy of correct answers.

We interpret the superior performance on LDT using words from Obama's speech as the result of phonological priming due to more frequent perceptual repetitions. However, we cannot exclude the additional contribution of other forms of priming such as semantic. Priming has been empirically demonstrated to improve comprehension of spoken language, particularly in adverse listening conditions (Sheldon, Pichora-Fuller, & Schneider, 2008). Therefore, we suggest that Obama's speech can be perceived as more clear and memorable due to lexical access

priming mechanisms.

In the last part of this paper, we explored the relation between the rhythmic melodic patterns of Obama's speech with music. By converting the prosodic information of writings into musical notation, we proved that Obama's speech can elicit emotional responses in the listeners, congruent with the contextual meaning of the speech. On the contrary, the writing of a schizophrenic subject elicited emotional responses incongruent with the context. This finding provides empirical support to Darwin's belief that "We must suppose that the rhythms and cadences of oratory are derived from previously developed musical powers" (Darwin, 1874).

The congruence of the context of speech with the emotional responses of volunteers listening to music, created with the prosodic information of Obama's speech, provides indirect evidence of emotional entrainment between the speaker and his audience. The opposite finding of incongruent emotional responses to the music extracted from the writing of a schizophrenic subject is in keeping with the known impaired expression of affective prosody in schizophrenia.

The material used in this study did not allow investigating other components of social entrainment. Our data do not allow speculating about any possible relation between lexical access priming and emotional entrainment, as well as their relative contribution to the salience of an oratorical speech.

Our findings bear upon the debate on the evolutionary basis of music by providing support to the view that "music is a biologically powerful human invention or transformative technology of the mind, that builds on diverse, pre-existing brain functions, rather than a trait that originated via processes of natural selection" (Patel, 2010). In fact, we show that the perception of melody, i.e., a series of single musical notes, is necessary to the generation and perception of emotional prosody of language, which, we argue, has a fundamental role in shaping the hierarchical structure of human social groups. Humans, through the analysis of both the semantic and prosodic components of language, were able to select individuals with those cognitive skills needed to lead their social groups.

Acknowledgements

References

Apfelbaum, K. S., Blumstein, S. E., & McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: Evidence for continuous cascading systems. *Psychon Bull Rev.*, *18*, 141-149. http://dx.doi.org/10.3758/s13423-010-0039-8

Ceccherini-Nelli, A., & Crow, T. J. (2003**)**. Disintegration of the components of language as the path to a revision of Bleuler's and Schneider's concepts of schizophrenia: Linguistic disturbances compared with first-rank symptoms in acute psychosis. *British Journal of Psychiatry*, *182*, 233-240. http://dx.doi.org/10.1192/bjp.182.3.233

Darwin, C. (1872). *The Expression of the Emotions in Man and Animals* (p. 83). London: Murray J. Ed. http://dx.doi.org/10.1037/10001-000

Darwin, C. (1874). *The Descent of Man and Selection in Relation to Sex* (2nd ed., p. 572). London: Murray J. Ed. http://dx.doi.org/10.5962/bhl.title.54341

Dawkins, R., & Krebs, J. R. (1978). *Animal signals: Information or manipulation. Behavioural Ecology: An Evolutionary Approach* (pp. 282-309). Oxford: Blackwell Scientific Publications.

Hatada, S., Sawada, K., Akamatsu, M., Doi, E., Minese, M., Yamashita, M., … Inoue, S. (2014). Impaired musical ability in people with schizophrenia. *J. Psychiatry Neurosci.*, *39*, 118-126. http://dx.doi.org/10.1503/jpn.120207

Hoekert, M., Kahn, R. S., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, *96*, 135-145. http://dx.doi.org/10.1016/j.schres.2007.07.023

Inquisit 3 [Computer software]. (2012). *Seattle, WA: Millisecond Software LLC*. Retrieved from http://www.millisecond.com

Jones, D., & Gregson, P. (2012). *The listening machine*. Retrieved from http://www.thelisteningmachine.org/about

Kahn, C. (2007). *P22 Music Text Composition Generator*. Retrieved from https://p22.com/musicfont/#toc

Kim, I., Mathon, C., & Boulakia, G. (2010). Rhetorical Prosody in French Courtroom Discourse. *Speech Prosody 2010-Fifth International Conference*. Retrieved from http://20.210-193-52.unknown.qala.com.sg/archive/sp2010/papers/ sp10_080.pdf

Lemmeti, S. (1999). *Master's Thesis: Review of Speech Synthesis Technology*. Retrieved from http://www.acoustics.hut.fi/publications/files/theses/lemmetty_mst/index.html

Levitan, R., Gravano, A., Wilson, L., Benus, S., Hirschberg, J., & Nenkova, A. (2012). Acoustic-Prosodic Entrainment and Social Behavior. *2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 11-19). Montreal, Canada. Retrieved from http://www.cs.columbia.edu/speech/PaperFiles/2012/N12-1002.pdf

Logan, H. M. (1976). The computer and the sound texture of poetry. *Language and Style*, *9*, 260-279.

Logan, H. M. (1985). Most by numbers judge the poet's song: Measuring sound effects in poetry. *Computers and the humanities*, *19*, 213-222. http://dx.doi.org/10.1007/BF02259574

Martınez, D., Lleida, E., Ortega, A., & Miguel, A. (2013). Prosodic features and formant modeling for an ivector-based language recognition system. *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference* (pp. 6847-6851). Vancouver, BC, Canada. http://dx.doi.org/10.1109/ICASSP.2013.6638988

Moskowitz, A., & Heim, G. (2011). Eugen Bleuler's Dementia Praecox or the Group of Schizophrenias (1911): A Centenary Appreciation and Reconsideration. *Schizophrenia Bulletin*, *37*, 471-479. http://dx.doi.org/10.1093/schbul/sbr016

Nash, J. F. (1950). Equilibrium points in N-person games. *PNAS*, *36*, 48-49. http://dx.doi.org/10.1073/pnas.36.1.48

Nash, J. F. (1958). Continuity of solutions of parabolic and elliptic equation. *American Journal of Mathematics*, 80, 931-954. http://dx.doi.org/10.2307/2372841

Nash, J. F. (1966). Analyticity of the Solutions of Implicit Function Problems With Analytic Data. *The Annals of Mathematics* (2nd Series), *84*, 345-355. http://dx.doi.org/10.2307/1970448

Night, S. (2009). Marching to the same beat: entrainment and persuasion. *Proceedings of the 2nd International Conference of Students of Systematic Musicology* (pp. 81-84). Ghent, Belgium.

Owings, D. H. & Morton, E. S. (1998). *Animal Vocal Communication: A new approach*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139167901

Oxford University Press. (n.d.). *Oxford Dictionaries*. Retrieved September 20, 2013, from http://oxforddictionaries.com

Parker, S. E. (2001). *Physics 4830 Notes Tues 11/13/01*. University of Colorado Boulder. Retrieved March 26, 2014, from http://www.colorado.edu/physics/phys4830/phys4830_fa01/lab/n1113.htm

Patel, A. (2010). Music, biological evolution, and the brain. In M. Bailer (Ed.), *Emerging Disciplines* (pp. 91-144). Houston, TX: Houston University Press.

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Percept*, *28*, 3-14. http://dx.doi.org/10.1525/mp.2010.28.1.3

Pinkham, A. E., Penn, D. L., Perkins, D. O., & Lieberman, J. (2003). Implications for the neural basis of social cognition for the study of schizophrenia. *American Journal of Psychiatry*, *160*, 815-824. http://dx.doi.org/10.1176/appi.ajp.160.5.815

Rouibah, A., Tiberghien, G., & Lupker, S. J. (1999). *Phonological and semantic priming: Evidence for task-independent effects. Mem Cognit.*, *27*, 422-437. http://dx.doi.org/10.3758/BF03211538

Sheldon, S., Pichora-Fuller, K. M., & Schneider, B. A. (2008). Priming and sentence context support listening to noise-decoded speech by younger and older adults. *J. Acoust. Soc.*, *123*, 489-99. http://dx.doi.org/10.1121/1.2783762

Suits, B. H. (1998). *Physics of Music Notes*. Retrieved from http://www.phy.mtu.edu/~suits/notefreqs.html

Wu, K.Y., Chao, C. W., Hung, C. I., Chen, Y. T., & Liang, S. F. (2013). Functional abnormalities in the cortical processing of sound complexity and musical consonance in schizophrenia: Evidence from an evoked potential stud. *BMC Psychiatry*, *13*, 158. http://dx.doi.org/10.1186/1471-244X-13-158

Yildirim, S., Bulut, M., Lee, C. M., Kazemzadeh, A., Deng, Z., Lee, S., … Busso, C. (2004). An Acoustic Study of Emotions Expressed in Speech. *Eighth International Conference on Spoken Language Processing*. Retrieved from http://graphics.nsm.uh.edu/wp-content/papers/2004/ICSLP04_acousticStudy.pdf

Zentner, M., & Eerola, T. (2009). Self-report based measures and models of musical emotion. In P. Juslin, & J. Sloboda (Eds.), *Handbook of Music and Emotion*. Oxford: Oxford University Press

Zuberbuhler, K. (2012). Primate Communication. *Nature Education Knowledge* (Vol. 3, p. 832).