

Do We Sing Like We Speak?

A Comparison of Question Intonation Patterns and Melodic Profile in Songs

No Author Given

No Institute Given

Abstract. This paper investigates whether intonation patterns in spoken language map to melodic profile patterns in music. The study focuses on polar (yes/no) and interrogative (wh) questions, and commonly found terminal intonation patterns for these kinds of questions in British English. We analyse the f0 of the main melody of 50 audio excerpts, from 49 different British pop songs by 47 distinct artists, for which the lyrics correspond to the desired questions types. Comparing the f0 of the main melody to the intonation patterns observed in authentic conversation, we test the null hypothesis that *there is no correlation between question intonation patterns in speech and melodic profiles in songs with question lyrics*. The data rejects the hypothesis, and we conclude that intonation patterns of questions in speech and in song melodies are correlated.

Keywords: music and speech, prosody, intonation, melodic profile, song lyrics, questions

1 Introduction

There are many different types of question that are commonly found within English language, and there are standard models of the common intonation patterns within the British English speech. This paper compares the fundamental frequency (f0) pattern of a song melody and the natural English speech intonation of same statement. The remit of this paper is to compare the intonation pattern of questions within music and natural speech in British English. Within speech, there is a natural intonation that is applied when someone utters a question and this paper intends to compare the intonations to the melody of questions as they appear in popular music.

Within this paper, the introduction will go on to discuss the motivation behind this work. Section 2 will present a literature review on common speech question intonation patterns and then go on to explain the work undertaken as part of this project and how the analysis will be performed. Section 3 will explain the data set that was collected as part of this project and how it was collected. The results of analyses of this data set will be presented in Section 4 and analysis of these results will be presented in Section 5. A conclusion will be drawn from this project and potential further work presented in Section 6.

1.1 Motivation

The studying of intonation and correlating speech and language would be interesting and useful work for music analysts, and assist understanding of motivations behind writing of musical melodies. It would also potentially provide useful insight into the psychology behind written music and musical structures. Further work, beyond the remit of this paper, could go on to explore universal comparison of intonation patterns within songs and music, which could introduce other applications of this work, including generative melody, based on a piece of text, allowing an automated rule based structure for allowing any written text to be turned into a musical performance. This could be beneficial as a potential memory technique, allowing an automated text to song generator could then be used to get a song stuck in your head and with recalling the melody the lyrics could be associated [1].

Language comparison and the way in which different language intonations and song intonation correlate could be an interesting extension of this work, with the potential for new language learning and intonation pattern learning techniques based on music and songs [2, 3].

The results from this paper could progress research in a wide range of potential directions for music and speech intonation. The possibilities for further research, focusing on other areas of natural speech intonation and how speech intonation related to musical melodies could prove a very useful tool for research in the music information retrieval (MIR) field, along with musicologists, music psychologists and linguists [1]. The potential for further research in comparison of intonation and music could also provide interesting exploration into international music, where the song is written in one country and language, but listened to in another, and its social impact on the language in the countries. Further studies into the musical call and response patterns and their correlation with natural language speech patterns could provide some valuable insight into the evolution of spoken language and social interactions.

2 Method

This section will present the method that will be implemented for this paper, with relation to the dataset collected as discussed in Section 3. A linguistic literature review will be presented, followed by a brief explanation of the data set to be collected, and then we shall present the analysis method used within this paper.

2.1 Speech Intonation for Comparison

Phonetics is the branch of linguistics that studies noises made as part of normal speech, including intonation patterns. Within this field, there are many studies into intonation relating to questions or different languages, but minimal studies with the aim of comparison with music. This section will discuss previous work

on speech intonation relating to questions and produce a phonetic model for comparison to our data set collected. There are several different types of question, and for the purpose of this paper we will be focusing on polar questions, which are questions with a yes or no response, and ‘wh’ questions, that are questions with an interrogative word, such as who or how [4]. For the purpose of this paper, the focus will be on the terminal intonation, the pattern of the end of the sentence, where the intonation patterns are most regular relating to questions [5, 6, 7]. Terminal phonetic patterns can be described in several manners, this paper will discuss intonation by the manner of change in the pitch.

Figure 1 demonstrates the phrasing and type of frequency pattern that would be expected. This model will be used throughout this paper.

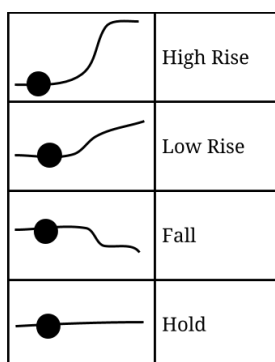


Fig. 1: Intonation Pattern Models

Grabe et al. [6] present the argument that both ‘wh’ and polar questions most commonly have a falling intonation pattern, though within polar questions the match was only 44% of questions matching this pattern. O’Connor and Arnold [8] present the argument for a low rise effect in polar questions and Cruttenden [9] suggest that a low rise or a fall are both suitable models for ‘wh’ questions. Huynh [10] compared intonation patterns within conversation and the textbook intonation patterns, and an excerpt is presented in Table 1.

For the purposes of this paper, songs will be grouped into the terminal intonation pattern groups, as either rise-fall, rise, fall, fall-rise, or hold. These groups will be compared against the textbook intonation patterns and general opinion of intonation.

2.2 Data Set

A data set will be produced as part of this paper. This data set will consist of a series of musical excerpts, cut from British English popular music. The tracks will be identified as either a polar question or a ‘wh’ question. These question

Table 1: Speech Intonation Patterns From [10]

Intonation Pattern	Authentic conversation	Textbook
Wh-question	Rising-falling Falling-rising	Rising-Falling
Yes/ No questions	Falling-Rising Terminal falling pitch	Falling-Rising

types will then be compared to a standard linguistic intonation model. Where question lyrics appear in a song, the fundamental frequency (f0) will be isolated and mapped.

2.3 Analysis

To analyse the results and compare the speech intonation to song lyric f0 track, a hypothesis is first produced:

Hypothesis *The intonation patterns of questions within speech and music are correlated.*

From these intonation patterns, abstracted from the audio data, each song will be classified dependant on the terminal intonation pattern that it follows. These results will be modelled and the mean of each intonation model will be plotted. These results can then be compared to standard textbook speech intonation patterns for the given question types. From this, it will be apparent how many of the song intonation patterns follow with the common speech patterns, and the results will be concluded.

3 Data Set

As part of this paper, a dataset was produced including collection, editing and analysis to produce a map of the fundamental frequency of the lyrics of each song. This section will discuss the collection and production of this data set.

3.1 Data Set Selection

A selection of fifty audio clips from forty-nine different songs and forty-seven different performers was selected. This data set was collected on the principal of the performer being British English. The songs were selected from a range of British Pop music with the year published over the past 60 years. The full data set is listed in the appendix [11]

The songs were selected had different question types in them, and each question type will be analysed within its groups to speech intonation patterns. The songs were classified, into two different categories, either polar questions or ‘wh’ questions.

3.2 Data Set Production

The data set was edited from audio tracks, and the appropriate section of audio was cut out using [12] software on an Apple Macbook Pro. From this the f0 was attained by the use of Tony software [13] with manual adjustment, by the author, to produce a f0 track of the sung melody. Figure 2 provides a screenshot of a selected track, in the Tony Annotation Environment, and Figure 3 presents the same annotation data plotted in Matlab.

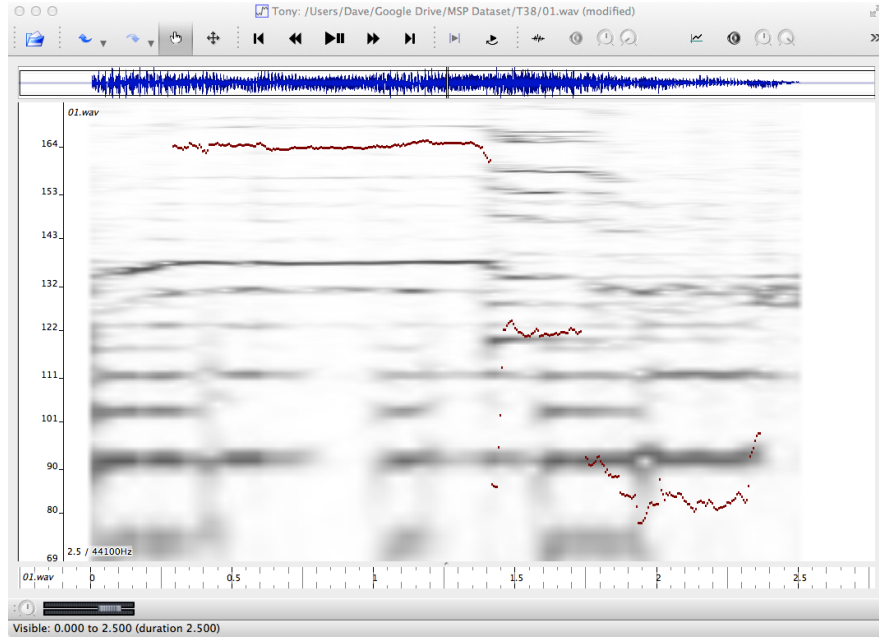


Fig. 2: A Screenshot of the Tony Annotation Environment

4 Results

In this section the results of comparing the f0 of the lyrics is compared to the expected intonation patterns, as occur in natural speech.

The results of the f0 were categorised as either a pitch rise-fall, a pitch fall or some other pitch pattern and the classification can be seen in Table 2 and 3. The f0 track of every song can be found in the appendix [11].

It can be seen that the majority of questions follow the terminal fall intonation pattern.

The mean f0 has been plotted for ‘wh’ questions with a falling and rise-fall intonation patterns in Figure 4a and Figure 4b respectively. All ‘wh’ question

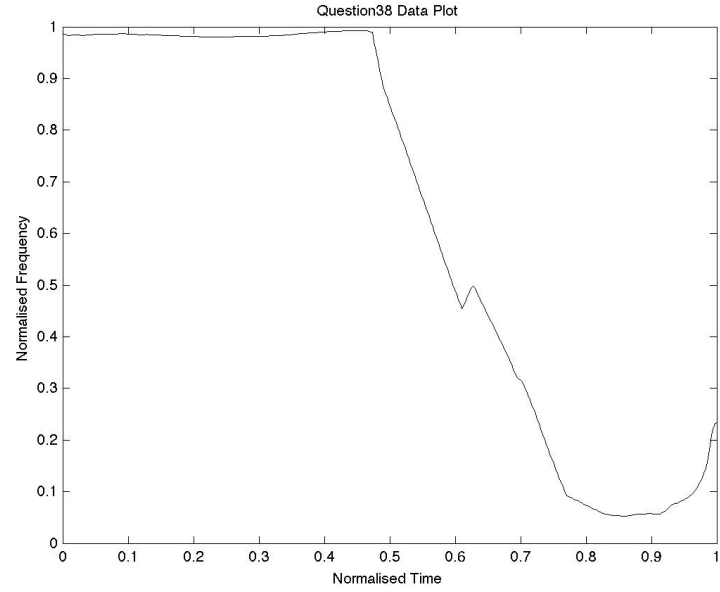


Fig. 3: The MATLAB plot of the annotation data, smoothed and normalised by both time and frequency.

Table 2: Intonation Pattern of Song Lyric Intonation Classification

Question Type	Rise-fall	Fall-Rise	Fall	Other	Total
Wh-question	11	6	5	3	25
Yes/ No questions	11	2	7	5	25
All Questions	22	8	12	8	50

Table 3: Intonation Pattern of Song Lyric Intonation Classification as Percentage

Question Type	Rise-fall	Fall-Rise	Fall	Other
Wh-question	44%	24%	20%	12%
Yes/ No questions	44%	8%	28%	20%
All Questions	44%	16%	24%	16%

intonation patterns with a falling intonation have been plotted in Figure 4c, with the mean f0 plotted in bold as a comparison, to demonstrate the spread of the data presented and likewise for the rising-falling intonations in Figure 4d.

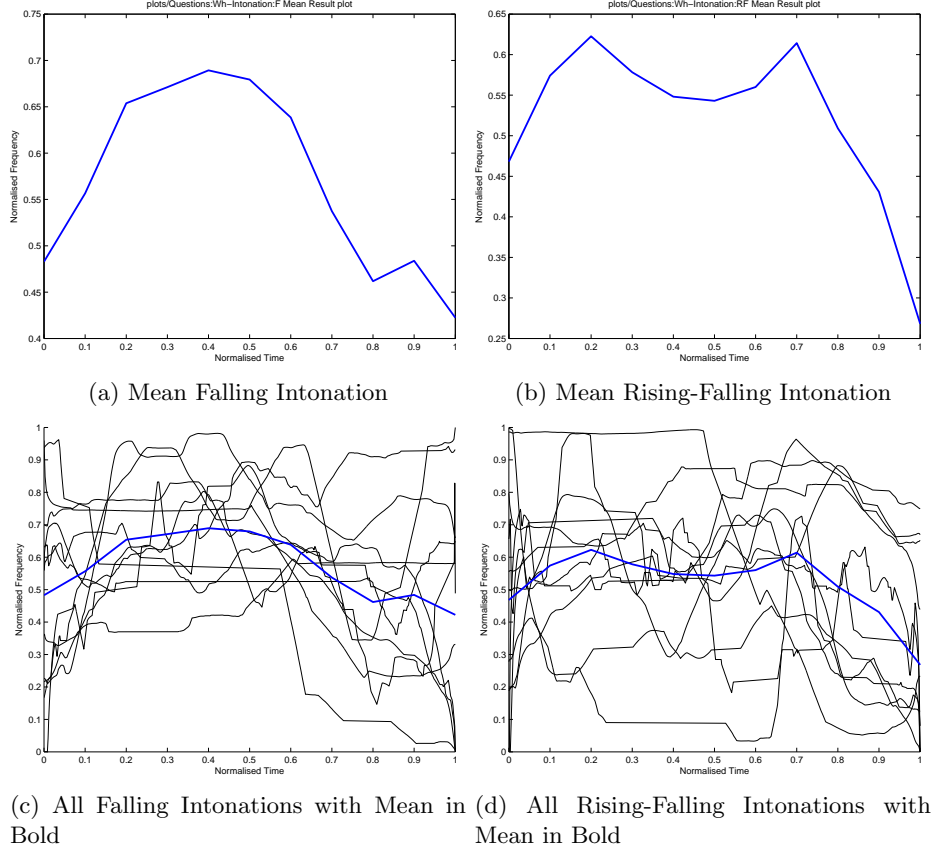


Fig. 4: ‘Wh’ Question Mean Frequency Plot

The mean f0 of polar questions with a falling and rise-fall intonation patterns are plotted in Figure 5a and Figure 5b respectfully.

5 Discussion

It can be seen from Table 3, that the majority of questions followed the falling intonation pattern. This correlates with the speech intonation patterns cited in Grabe et al. [6], who suggested that questions should all have a terminal falling intonation pattern. Both rise-fall and fall intonations have a terminal fall intonation pattern, and thus 64% of interrogative questions and 72% of

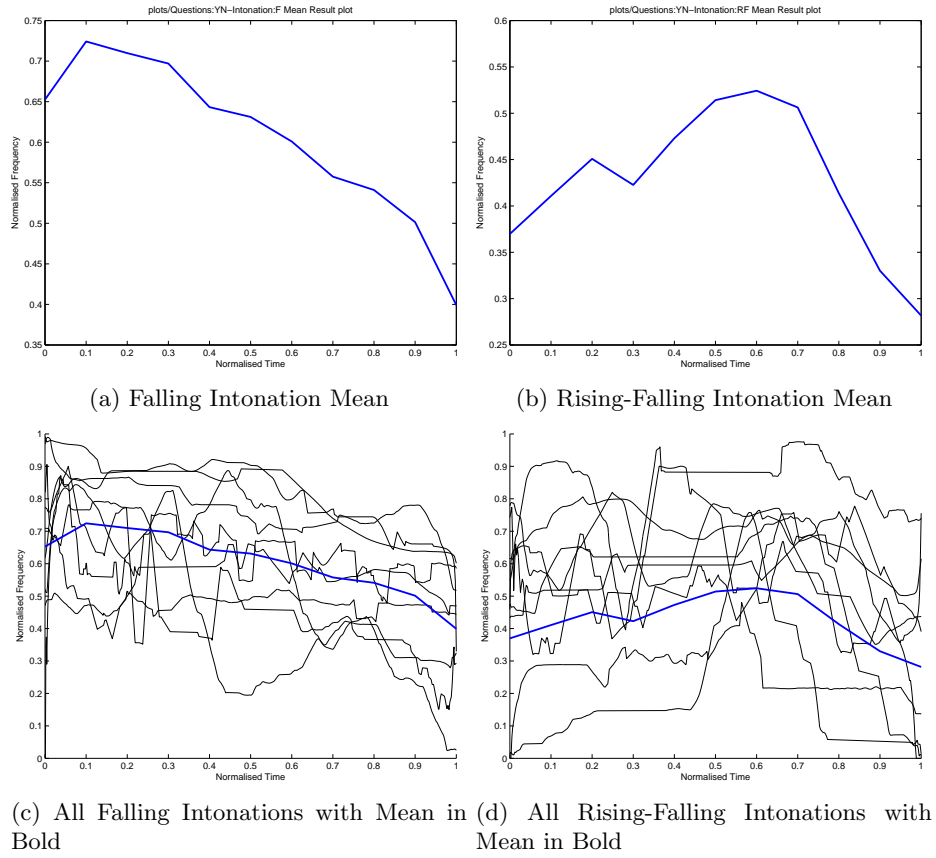


Fig. 5: Polar Question Mean Frequency Plot

polar questions have a terminal falling intonation. These results also concur with authentic conversations intonation from Huynh [10] and standard intonation patterns as proposed by Cruttenden [9] regarding the terminal intonation pitch as a falling pattern. Thus our hypothesis of *The intonation patterns of questions within speech and music are correlated.* has been agreed.

These results are statistically analysed to verify correlations. A set of example data is taken from Grabe et al. [6], where speech intonation patterns for ‘Wh’ and polar question intonation patterns are given. This provides us with a set of comparative data, which can be used to calculate the correlation between our gathered dataset and some other real world data, in an attempt to prove our hypothesis.

Table 4: Intonation Pattern for Questions from Grabe et al. [6]

Intonation	‘Wh’ Questions			Polar Questions			Total
	Cambridge	Newcastle	Belfast	Cambridge	Newcastle	Belfast	
RF	0	18	0	0	0	0	18
FR	19	24	90	30	32	24	219
F	57	42	0	36	25	0	160
R	32	12	12	18	0	66	140
H	0	0	0	0	0	0	0
Total	108	96	102	84	57	90	537

Table 5: χ^2 Value and Probability of Null Hypothesis for Different Question Types

Question Type	χ^2	P(Null Hypothesis)
‘Wh’ Questions	55.17	0
Polar Questions	119.2	0

Question results used from Grabe et al. [6] are presented in Table 4. This data was used to verify the following null hypothesis:

Null Hypothesis *There is no correlation between speech intonation patterns and song lyric intonation patterns.*

The χ^2 value was calculated, as a test of independence between the gather dataset and the Grabe et al. [6] dataset and the results are presented in table 5. It can be seen from table 5, the null hypothesis has been disproved, with over

99% confidence and therefore there is a correlation between song lyric intonation patterns where questions are present and in speech, in British English.

There are strong and statistically significant correlations between the song lyric dataset and the Grabe question dataset. It is the conclusion of this work that speech intonation patterns correlate strongly with the pitch of melodies within popular music containing questions, and these results have been justified as statistically significant.

6 Conclusion

This paper presents how a dataset of 50 excerpts of British Pop music was produced, where every excerpt consists of a song lyric containing a question. These song lyric melodic profiles were annotated and classified, using linguistic techniques, with regards to their terminal intonation pattern. These terminal intonation patterns were then compared to textbook speech intonation patterns and a summarised dataset taken from Grabe et al. [6], to verify the null hypothesis *There is no correlation between speech intonation patterns and song lyric intonation patterns*. The null hypothesis was rejected and we conclude that there are correlations between the intonation patterns of questions in speech and song melodies.

7 Further Work

This paper has revealed a range of potential further work in this field. It is possible to extend this work by focusing on dialects of specific song writers, to clarify that song writers do indeed follow their specific dialects intonation patterns. As discussed in Grabe et al. [6], there are a range of dialects, each with their own intonation patterns that follow, and so further detailed study in this area could be performed. This could also be extended to clarify if there are any differences in the linguistic patterns based if the melody is written and lyrics fitted to the melody, or if the song writing process attempts to fit a melody to the lyrics.

This work could be extended to review a larger range of languages and dialects, or to review other musical genres other than British Pop music. This may allow further work to be carried out with the use of symbolic data, rather than requiring manual annotation of songs, which could easily allow for a significant increase in dataset with access to the musical scores and song lyrics.

References

- [1] Wiering, F.: Can humans benefit from music information retrieval? In: Adaptive Multimedia Retrieval: User, Context, and Feedback, pp. 82–94. Springer (2007)
- [2] Nichols, E., Morris, D., Basu, S., Raphael, C.: Relationships between lyrics and melody in popular music. In: ISMIR. pp. 471–476 (2009)
- [3] Zatorre, R.J., Baum, S.R.: Musical melody and speech intonation: singing a different tune. PLoS biology 10(7), e1001372 (2012)
- [4] De Pijper, J.R.: Modelling British English Intonation, vol. 3. Walter de Gruyter (1983)
- [5] Eady, S.J.: Differences in the f0 patterns of speech: Tone language versus stress language. Language and Speech 25(1), 29–42 (1982)
- [6] Grabe, E., Kochanski, G., Coleman, J.: The intonation of native accent varieties in the British Isles: potential for miscommunication. English pronunciation models: a changing scene pp. 311–338 (2005)
- [7] Hirst, D.: Intonation in British English. Intonation Systems. A Survey of Twenty Languages 56, 77 (1998)
- [8] O’Connor, J.D., Arnold, G.F.: 1973, Intonation of colloquial English (1961)
- [9] Cruttenden, A.: On the so-called grammatical function of intonation. Phonetica 21(3), 182–192 (1970)
- [10] Huynh, L.T.: Question intonation patterns in a real-life conversation and in textbook dialogs. TESOL Working Paper Series 10, 83–92. (2012)
- [11] Author: Modelling of intonation from speech to music appendix (January 2015), <https://code.soundsoftware.ac.uk/projects/modelling-of-intonation-from-speech-to-music-appendix>
- [12] Logic: Logic Pro X, version 10.0.6. Apple Inc., Cupertino, CA, United States of America (2014)
- [13] Mauch, M., Dixon, S.: Tony, a tool for melody annotation, version 0.4. Queen Mary University London (2014)
- [14] Hermes, D.J., Van Gestel, J.C.: The frequency scale of speech intonation. The Journal of the Acoustical Society of America 90(1), 97–102 (1991)
- [15] Grabe, E., Post, B.: Intonational variation in the British Isles. In: Speech Prosody 2002, International Conference (2002)
- [16] Mauch, M., Dixon, S.: pYIN: A fundamental frequency estimator using probabilistic threshold distributions. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2014) (2014), under review
- [17] Bech, S., Zacharov, N.: Perceptual audio evaluation-Theory, method and application. John Wiley & Sons (2007)
- [18] Devaney, J., Mandel, M.I., Fujinaga, I.: A study of intonation in three-part singing using the automatic music performance analysis and comparison toolkit (AMPACT). In: ISMIR. pp. 511–516 (2012)