

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 exist many different types of questions that are commonly found within the English language, and there exist standard models of the common intonation patterns for these questions types within British English speech. We are interested in investigating whether these intonation patterns translate to melodic patterns when the words are put to song. The remit of this paper is to compare the intonation pattern of questions within music and natural speech in British English. In particular, this paper aims to compare the intonation patterns applied when someone utters a question to the melody of the same types of questions when they appear in popular music. More specifically, this paper compares the fundamental frequency (f0) pattern of the song melody and the natural English speech intonation of the same statement.

Phonetics is the branch of linguistics that studies the sounds made as part of normal speech, including intonation patterns. Intonation, the variation of pitch, in spoken language is frequently used to signal the ends of questions. Within the field of phonetics, many studies exist pertaining to intonation of questions or of different languages, but there is only limited work comparing intonation patterns to music. We describe the most relevant research here.

Zatorre and Baum [1] evaluates the differences between human perception of music and speech, focusing on interval and frequency patterns. Ross et al. [2] explores musical intervals and standard western music chord structure, and correlates this with speech intervals. List [3] identifies the fundamental differences between music and speech, discussing classification differences between the two sets of sound. Fonagy and Magdics [4] attempts to classify both music and speech data sets into emotional categories based on melodic patterns.

[EC: NEED TO ALSO CITE THE ANI PATEL PAPER ON MUSIC AND SPEECH NPVI (MEASURING DURATION CONTRAST) AND CV (MELODIC INTERVAL CONTRAST) COMPARISONS. AND FOLLOW UP WORK BY HURON. SEE MusicalPatois LECTURE NOTES.]

We take existing efforts a step further to focus specifically on spoken and sung questions. There are several different types of questions, with varying intonation patterns in British English. 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, which are questions with an interrogative word, such as who or how [5]. For this study, we will focus on the terminal intonation, the pattern of the end of the sentence, where question intonation patterns are most regular [6, 7, 8]. Comparing these terminal intonation patterns to that found in song melodies corresponding to the question types of interest, we ascertain whether speech intonation and song melody patterns correlate directly.

In order to conduct the study, we created a dataset consisting of excerpts from British English popular music corresponding to lyrics determined as either polar questions or ‘wh’ questions. Prior work on speech intonation of questions provides a phonetic model for comparison to our sung melody data set, and we test the hypothesis that intonation patterns of spoken questions and sung melodic profiles corresponding to questions are correlated.

The study of how speech intonation and song melodies inter-relate benefits the understanding of the art of composing melodies for songs. It provides useful insight into the psychology of music writing and the design of melodic structures. Applications include the generating of natural sounding melodies (any written text could be turned into a musical utterance using the rules learned), language acquisition (intonation pattern learning using music and songs) [17, 1], and culturally-sensitive lyric translations. The knowledge gained could also benefit music retrieval, for example, allowing earworms without words to be more readily associated with their lyrics [9].

The remainder of this paper is organised as follows: Section 2 presents the work undertaken and explains how the data set was created and the analysis performed; the results of the analyses and discussions of the results are presented in Section 3; conclusions are drawn and further work presented in Section 4.

2 Method

This section presents the method implemented for this paper, with relation to the dataset collected as discussed in Section 2.2.

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 Basis for Comparison: Speech Intonation

Terminal phonetic patterns of questions can be represented in various ways; here, we focus on pitch change. The commonly used “tadpole diagram” representation is shown in Figure 1. The line drawings indicating the shape of the fundamental frequency of the intonation pattern; the head indicates the point of stress. These models of terminal phonetic patterns will be the ones referenced throughout the paper.

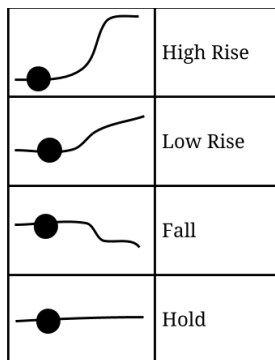


Fig. 1: Intonation pattern models

Next, we discuss some intonation patterns that have been theorised about and observed in authentic speech. Grabe et al. [7] present the argument that both ‘wh’ and polar questions most commonly have a falling intonation pattern, and found that only 44% of questions matched this pattern in spoken polar questions. O’Connor and Arnold [10] present the argument for a low rise effect in polar questions, and Cruttenden [11] suggest that a low rise or a fall are both suitable models for ‘wh’ questions. Huynh [12] compared intonation patterns in authentic conversation and textbook intonation patterns; a part of their results is reproduced in Table 1.

For the purposes of this paper, songs will grouped into the terminal intonation pattern groups: rise-fall, rise, fall, fall-rise, or hold. These groups will be compared against the textbook intonation patterns and **intonation patterns found by researchers in authentic conversation.(?)**

Table 1: Speech intonation patterns (reproduced from [12])

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

2.2 Data Set

For this study, a dataset was produced, including collection, editing and analysis of the data 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.

A selection of fifty audio clips from forty-nine different songs and forty-seven different performers was selected. A defining characteristic of this data set was that the performer **QUESTION: WAS THE PERFORMER ALSO THE COMPOSER?** had to be British English. The songs were selected from a range of British Pop music published over the past 60 years. A listing of the full data set is made available in the public Soundsoftware repository [13][**EC: NEED TO ANONYMISE THIS FOR BLIND REVIEW**] and reproduced in **Appendix Y in this paper** [**EC: PUT THE LIST OF SONGS AT THE END OF THIS PAPER IN AN APPENDIX**].

The songs were selected for the different question types in them—polar questions and ‘wh’ questions. The data set was edited from audio tracks; the appropriate section of audio was cut out using the Logic [14] software on an Apple Macbook Pro. The fundamental frequency (f_0) of the melody of each question was then obtained from the recording excerpt using the Tony [15] software (which uses the pYIN algorithm from [16]). A range of proposed f_0 tracks are proposed by the TONY software, and the correct f_0 representing the sung melody is selected manually. Figure 2 shows a selected track in the TONY annotation environment. The output is a series of non-linearly spaced data points. The data is then smoothed with a moving average window of 50 TONY output samples, and plotted in MATLAB; an example can be seen in Figure 3.

2.3 Analysis

The sung melodic profile patterns abstracted from the audio data are classified into one of four terminal intonation patterns: Rise-Fall, Fall-Rise, Fall, Others. We test the hypothesis that the intonation patterns of questions within speech and music are correlated. The dataset is analysed by comparing the song lyric f_0 tracks to a dataset of speech intonation patterns for questions. The speech dataset we use for the comparison consists of over 500 question intonation patterns for standard British English speech from a range of dialects, and is taken from [7]. From this, we determine whether the melodic profile patterns follow commonly observed speech patterns, and draw conclusions from the results.

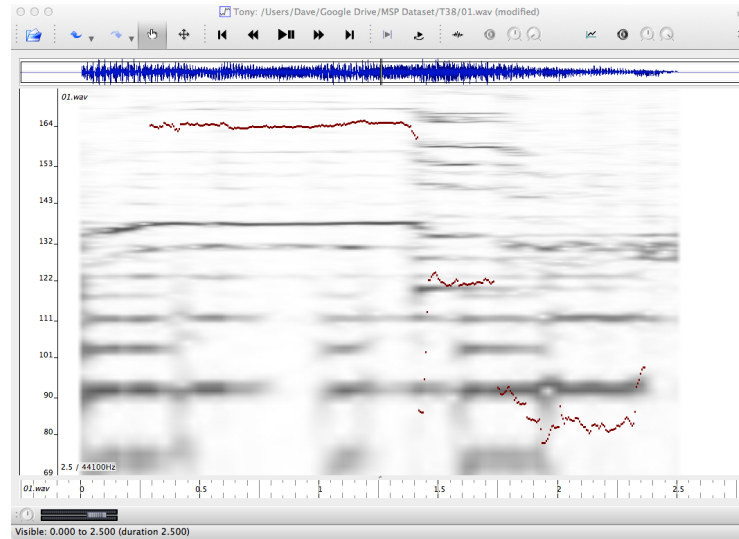


Fig. 2: A Screenshot of the TONY Annotation Environment corresponding to the words, “Who are you?” from the song, “Who are you?” by The Who

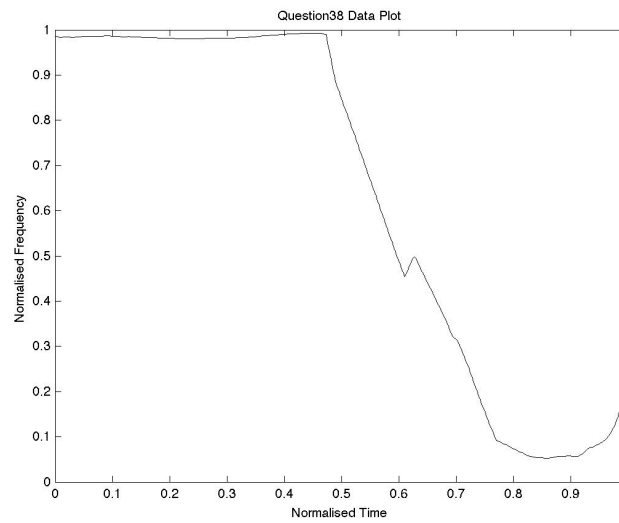


Fig. 3: A MATLAB plot of the TONY annotation data, smoothed and normalised by both time and frequency, corresponding to the data shown in Figure 2

3 Results and Discussion

In this section, we present the results of comparing the f_0 of the sung melodies to the expected speech intonation patterns, as they occur in natural speech, for the question types represented by the song lyrics. The f_0 track of every song has been made publicly available via the SoundSoftware repository [13] **EC: NEED TO ANONYMISE THIS FOR BLIND REVIEW**. The melody f_0 patterns extracted are categorised as into Rise-Fall, Fall-Rise, Fall, and Other patterns by inspection.

Figure 5 shows the f_0 frequencies of the sung melodies extracted in the manner described in Section 2.2 for the song excerpts corresponding to ‘wh’ questions. Figure 4a and Figure 4b show the melody profile patterns that were classified into Falling and Rise-Fall types, respectively. All ‘wh’ question sung melody profile patterns classified as Falling have been plotted in Figure 4c, with the mean f_0 plotted in bold for comparison, to demonstrate the spread of the data, and likewise for the Rise-Fall melodic profiles in Figure 4d.

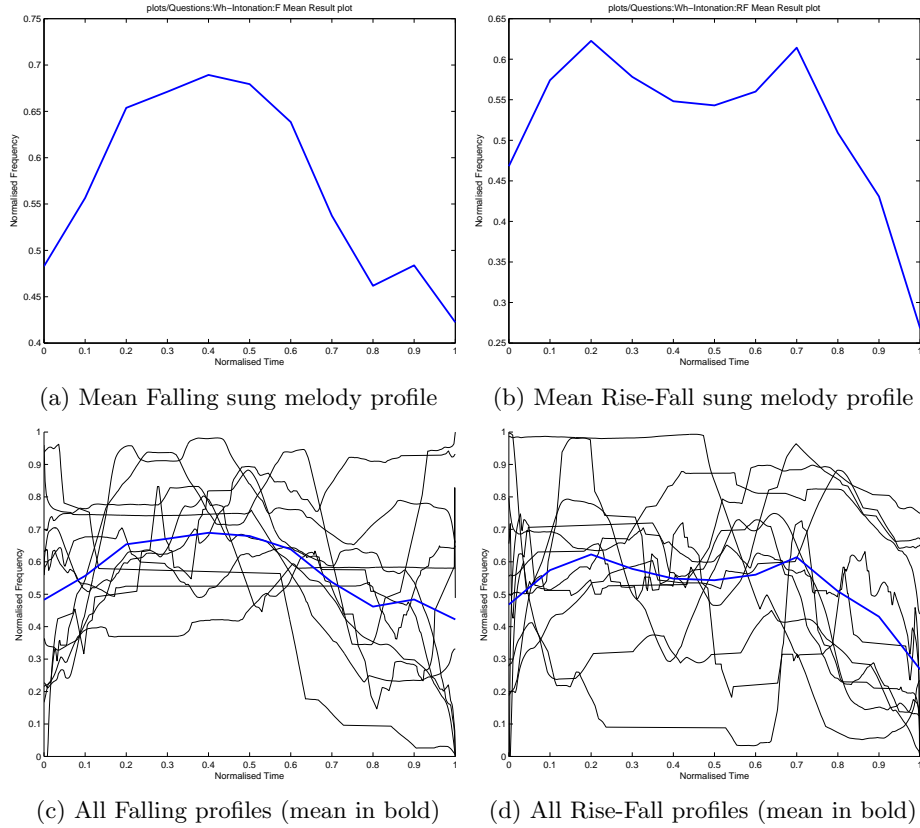


Fig. 4: Frequency (f_0) plots for ‘wh’ question sung melody profiles

The f_0 patterns of the sung melodies for polar question lyrics are plotted in Figure 5. As with the ‘wh’ questions, the mean sung melody profile is shown in Figure 5a and Figure 5b for Falling and Rise-Fall patterns, respectively, and all melody profiles are shown in Figure 5c and Figure 5d, respectively.

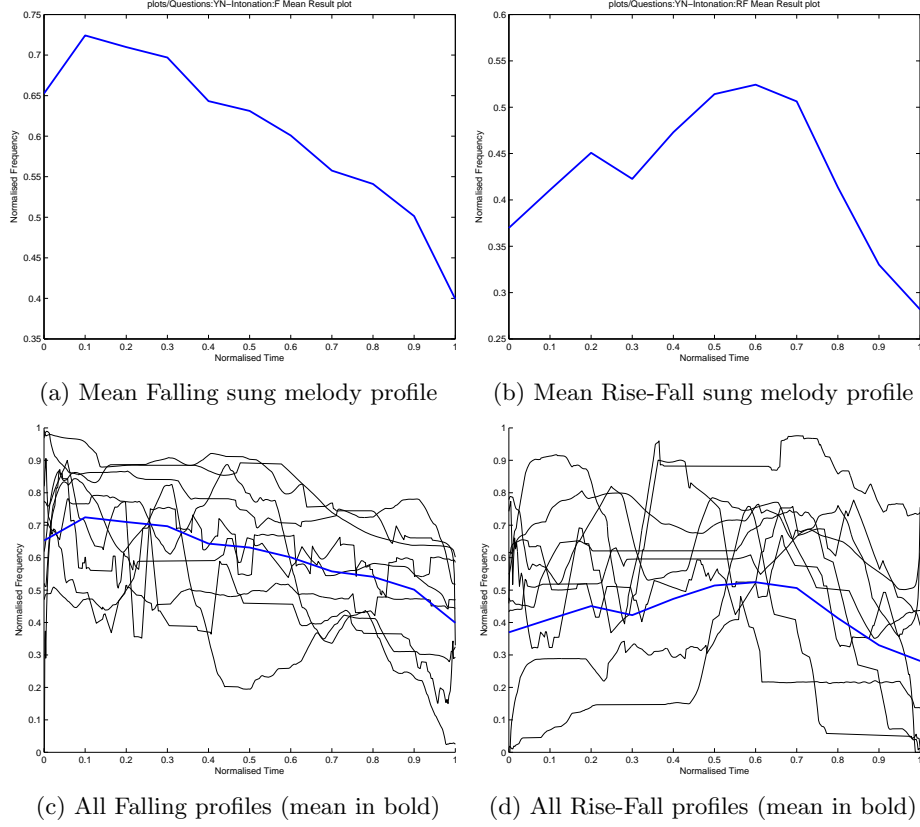


Fig. 5: Frequency (f_0) plots for polar question sung melody profiles

The numeric results of the classification are shown in Table 2 and Table 3. It can be seen that the majority of the terminal melodic profiles of questions follow the Falling intonation pattern. This correlates with the speech intonation patterns cited in Grabe et al. [7], in which it was suggested that questions should all have a terminal Falling intonation pattern. Both Rise-Fall and Falling melodic profiles exhibit a terminal falling pattern, and thus 64% of interrogative questions and 72% of polar questions had a terminal falling sung melody pattern.

These results also concur with intonation patterns observed in authentic conversations as described in Huynh [12], and with standard intonation patterns as proposed by Cruttenden [11] regarding the terminal falling intonation pattern.

Table 2: Classification of sung melodic profiles corresponding to questions in lyrics sorted by question types (absolute numbers)

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: Classification of sung melodic profiles corresponding to questions in lyrics sorted by question types (percentages)

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%

Thus, at first blush, our hypothesis that the intonation patterns of questions within speech and music are correlated appears to be supported by the data.

We employ statistical analysis to verify the observed correlations. We use as the template speech patterns of ‘wh’ and polar question a set of example data from Grabe et al. [7], reproduced in Table 4. This data provides us with a basis for comparison to test the null hypothesis that there is no correlation between speech intonation patterns and song lyric intonation patterns.

Table 4: Speech intonation patterns for different questions (reproduced from [7])

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

The χ^2 value was calculated to test the independence between the observed patterns in the melodic profiles dataset and that of Grabe et al. [7], and the results are presented in Table 5. The null hypothesis is rejected with over 99% confidence, and therefore we conclude that dependencies exist between song lyric melodic profile patterns corresponding to questions in British Popular music and terminal question intonation patterns in British English speech.

Table 5: χ^2 values and p-values for different question types

Question Type	χ^2	p-value
‘Wh’ Questions	55.17	0 [EC: PLEASE SHOW AT LEAST 4 DECIMAL POINTS, E.G. 0.0001]
Polar Questions	119.2	0

4 Conclusion

We have presented the creation of a sung melody dataset consisting of 50 excerpts of British Popular music corresponding to questions in the lyrics. These song lyric melodic profiles were annotated and classified with regards to their terminal intonation patterns. These terminal patterns were then compared to textbook speech intonation patterns and to a summary table from Grabe et al. [7] to test the null hypothesis that there is no correlation between speech intonation patterns and song lyric intonation patterns. The null hypothesis was rejected with high confidence and we conclude that correlations exist between the intonation patterns of questions in speech and the melodic profiles of sung melodies.

Our study focused on sung melody extracted from audio recordings. Similar studies could be carried out using symbolic data. This work could be extended to a larger range of languages and dialects, or to genres other than British Pop music. As discussed in Grabe et al. [7], there exists a range of dialects even of British English alone, each with their own intonation patterns. A straightforward extension of this work could examine similar dependencies for data sorted by the songwriters’ dialects to determine if composed melodies follow the specific dialect intonation patterns of the author.

References

- [1] Zatorre, R.J., Baum, S.R.: Musical melody and speech intonation: singing a different tune. *PLoS biology* 10(7), e1001372 (2012)
- [2] Ross, D., Choi, J., Purves, D.: Musical intervals in speech. *Proceedings of the National Academy of Sciences* 104(23), 9852–9857 (2007)
- [3] List, G.: The boundaries of speech and song. *Ethnomusicology* 7(1), 1–16 (1963)
- [4] Fonagy, I., Magdics, K.: Emotional patterns in intonation and music. *STUF-Language Typology and Universals* 16(1-4), 293–326 (1963)
- [5] De Pijper, J.R.: *Modelling British English Intonation*, vol. 3. Walter de Gruyter (1983)
- [6] Eady, S.J.: Differences in the f_0 patterns of speech: Tone language versus stress language. *Language and Speech* 25(1), 29–42 (1982)
- [7] 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)
- [8] Hirst, D.: Intonation in British English. *Intonation Systems. A Survey of Twenty Languages* 56, 77 (1998)
- [9] Wiering, F.: Can humans benefit from music information retrieval? In: *Adaptive Multimedia Retrieval: User, Context, and Feedback*, pp. 82–94. Springer (2007)
- [10] O’Connor, J.D., Arnold, G.F.: 1973, *Intonation of colloquial English* (1961)
- [11] Cruttenden, A.: On the so-called grammatical function of intonation. *Phonetica* 21(3), 182–192 (1970)
- [12] Huynh, L.T.: Question intonation patterns in a real-life conversation and in textbook dialogs. *TESOL Working Paper Series* 10, 83–92. (2012)
- [13] Author: Do we sing like we speak appendix (January 2015), <https://code.soundsoftware.ac.uk/projects/do-we-sing-like-we-speak>
- [14] Logic: *Logic Pro X*, version 10.0.6. Apple Inc., Cupertino, CA, United States of America (2014)
- [15] Mauch, M., Dixon, S.: *Tony*, a tool for melody annotation, version 0.4. Queen Mary University London (2014)
- [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] Nichols, E., Morris, D., Basu, S., Raphael, C.: Relationships between lyrics and melody in popular music. In: *ISMIR*. pp. 471–476 (2009)
- [18] 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)
- [19] Grabe, E., Post, B.: Intonational variation in the British Isles. In: *Speech Prosody 2002, International Conference* (2002)