

Harmonic Words, Narrative Chords: Textual Sonification Using Part of Speech

Jiin Ko
Artist, Researcher
jiinkollective@gmail.com

Youngjun Choi
TX Creative Media Lab
KAIST
youngjun.choi@kaist.ac.kr

Jinjoon Lee*
TX Creative Media Lab
KAIST
jinjoon.lee@kaist.ac.kr

ABSTRACT

This research aims to identify the commonalities between words (language) and chords (music), exploring if text itself can be musical and harmonic. Rather than letters or phonemes, we focus on words and their relationship between each other and have realized texts as music by using part of speech as chord progression on an automated algorithm. Other elements such as length and sentiment information inherent in text were also assigned to corresponding musical elements. For this research, a full-text of George Orwell's *Animal Farm* was turned into a three-hour harmonic piece.

Author Keywords

Textual sonification, part of speech, chord progression, algorithmic composition, text analysis

CCS Concepts

•Applied computing → Sound and music computing; Performing arts;

1. INTRODUCTION

“... while reading appears to be silent, in fact it is like a raging wave. This is what we call the harmony of reading.” As the Chinese author Yu Hua wrote [11], when static sentences in a book encounter a reader, the unseen harmony is brought to life. In addition to the cognitive processes involved in reading, the cause that creates the sound lies in the structure of language and its musical aspect existing in it.

Language and music, in this context referring to music with traditional structure, are analogous in that they are both divided by phrases causing intonation and entail sequences of sounds [14]. This paper explores the harmonic aspect revealed in part of speech in text and its application in chords. Part of speech was used to emphasize the relationship and transition of words and to express the ‘process’

*Corresponding author.

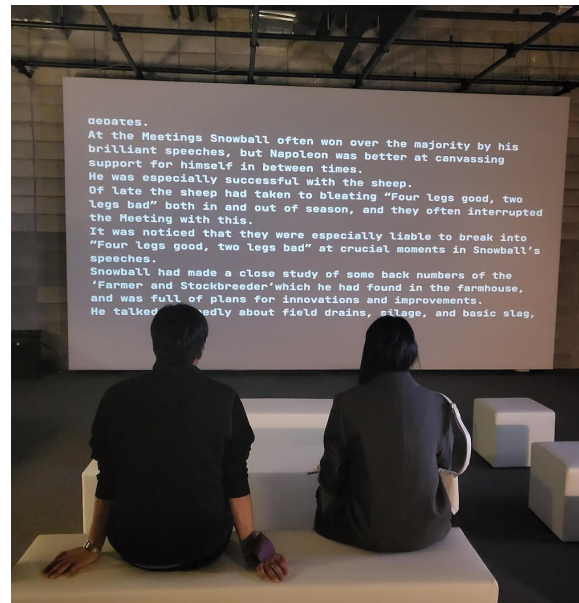


Figure 1: *Animal Farm Sings* displayed at Platform L, Seoul, 2022.

of reading because it is a dynamic and complex movement unfolding through time [8] as music is. For this research, the text of George Orwell's *Animal Farm* [17] was analyzed, and with the aid of an automated algorithm, it was recreated as a piece of music. Figure 1 shows how it was displayed in an exhibition. The audience could follow the words appearing on the screen while listening to its corresponding chords.

2. RELATED WORK

Barlow [6] mapped 25 letters from Beckett's Ping to musical keys using predefined rules. Text messages were turned into meaningful melodies [5], and Chinese poetry was transformed to music to reinforce synaesthesia [12]. There are programs that generate music from text as well such as MusicLM [3], Melobyte [2] and Singling [16]. While many approaches deal with phonological aspects of text-to-sound conversion, our project explores the harmonic potential of text and its transformation into musical compositions.

3. SIMILARITIES OF WORDS AND CHORDS

Lerdahl and Jackendoff's research [15] is seminal in exploring the relationship between language and music. In addition to the cognitive affinity they share, it has been argued that there are similarities in sound, structure, and general properties [10]. What we focus on here is the structure level



Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'24, 4 Sep–6 Sep, 2024, Online and Utrecht, Netherlands.

and some general properties. Sequence of words creates a sentence as sequence of chords creates harmony. Such parallel mechanism can be further illustrated by the following two propositions.

3.1 Structural Similarity

Words are classified by the functional trait called ‘part of speech (POS)’, one of the grammatical groups into which words are divided depending on their use [23]. It shows what function and role a word has in a sentence as a chord does in music. The order in which words are connected in a sentence is generally fixed and follows grammatical rules, with exceptions for special cases [9]. There are also rules in connecting chords to form phrases in music, especially in traditional music. Dmitri Tymoczko talks about harmonic syntax and the resemblance Western classical music bears to language. As in English where there is a subject-verb-object order, there is tonics – subdominants – dominants order in Western classical music [22].

3.2 Transitional Similarity

Narrowed down from language and music, words and chords share transitional similarity. Both language and music are subject to time, with their sequences unfolding temporally [18]. This feature leads us inevitably to the transition of words and chords moving from the current element to the next one. This quality is explained by Markov chain which is widely used for algorithmic composition for its function that calculates the probability of the element that comes after. This feature is very useful for natural language processing (NLP) as well. If we know what the current word is, we can predict or guess which word comes next [4]. It is also used for POS-tagging provided by Natural Language Tool Kit (NLTK). POS-tagging is the process of labelling POS to each word accordingly [7].

4. TEXTUAL SONIFICATION

For this project, literary works were chosen for their rule-based structure as this sonification process relies on grammatical rules, not on utterances [14]. Therefore, any types of refined text that can be divided into POS can be used for this sonification. To turn words to chords, the following process took place. Figure 2 shows a hierarchical overview of entire data processing.

4.1 Getting POS from Text

First, misspellings and typographical errors were corrected in Python and the text was inputted to Max to be parsed and processed for music and video. The values necessary for music were routed to Ableton Live. We used the universal tagset from NLTK which provides the generic terms of word classes in twelve categories [19] including punctuation marks and catch-all categories. Some tags were manually corrected to address limitations of the automatic tagging.

4.2 Realizing as Music

For *Animal Farm*, orchestral instruments were arranged to give sarcastic yet grand feel. This requires one’s knowledge on the text as one should set the general genre of the music that fits the story and mood of the text. Duration of a note is decided by the word length and voice leading of a triad (or a tetrad in case of subV7) is decided randomly. Speaker information played a crucial role in allocating instruments.

Part of Speech	Major	Harmonic minor
NOUN (nouns)	I	i
ADP (prepositions), PRT (particles)	ii	ii
PRON (pronouns), PROPER (proper nouns)	iii	bIII+
ADV (adverbs)	IV	iv
VERB (verbs)	V	V
ADJ (adjectives)	vi	bVI
CONJ (conjunctions)	vii°	vii°
DET (determiners and articles)	SubV7(Iib7)	SubV7 (Iib7)

Table 1: POS and corresponding chords

4.2.1 Turning POS Into Chords

POS tags were then turned into chords based on the rules established. Table 1 shows English POS and their corresponding chords. This process of matching words and chords involves subjective choices based on their perceived shared characteristics. Further research could explore alternative mappings and validate these choices. For minor scale, harmonic minor was used because it has more distinctive and audible tension than the other two minor scales. Table 2 shows the chords and keys applied in *The Seven Commandments in Animal Farm* and below are the rules applied.

- **I – Noun:** I or ‘1 chord’ is the tonic that determines the first note of a key acting as a basis. Nouns are fundamental in a sentence playing a similar role.
- **ii – Preposition and particle:** ii or ‘2m chord’ tends to resolve to the tonic. Prepositions combined with nouns form prepositional phrases.
- **iii - Proper noun and pronoun:** Proper nouns and pronouns act as substitutes for nouns as iii or ‘3m chord’ is to the tonic.
- **IV – Adverb:** IV or ‘4 chord’ is the subdominant that has a tendency to move to the dominant. Adverbs modify verbs [21] which is classified as the dominant here.
- **V – Verb:** V or ‘5 chord’ is the dominant. It requires the tonic for its resolution. Verbs are similar in a way that they take nouns as object to complete a sentence.
- **vi - Adjective:** vi or ‘6m chord’ is a relative minor of I. Adjectives, being related to nouns, modify or describe them [21].
- **vii° - Conjunctions:** The unstable feeling of vii° or ‘7dim chord’ matches the nature of conjunction that makes us look forward to the next word.
- **Iib7 (SubV7) – Iib7 or SubV7** is used as a substitute for V that goes to I. The fact that it should be resolved to the tonic is similar in a way determiners and articles should be followed by nouns.
- **Punctuation – Pause:** Punctuations such as period, exclamation mark, or question mark at the end of a sentence act as a pause in music.
- **Numbers and X – Nil:** Nothing was assigned to numbers and other catch-alls because they would interrupt the development of music.

4.2.2 Speaker Information

There are nine main characters including the narrator and instruments were allocated to each character based on their characteristics and traits. When a certain speaker talks, the assigned instrument to the speaker plays solo. For example, Napoleon the dictator was assigned a tuba and Squealer the flatterer a piccolo. A number of instruments play at the same time in accordance with how many sentences there are in a paragraph that the ensemble gets bigger as the paragraph gets longer.

4.2.3 Sentiment Information

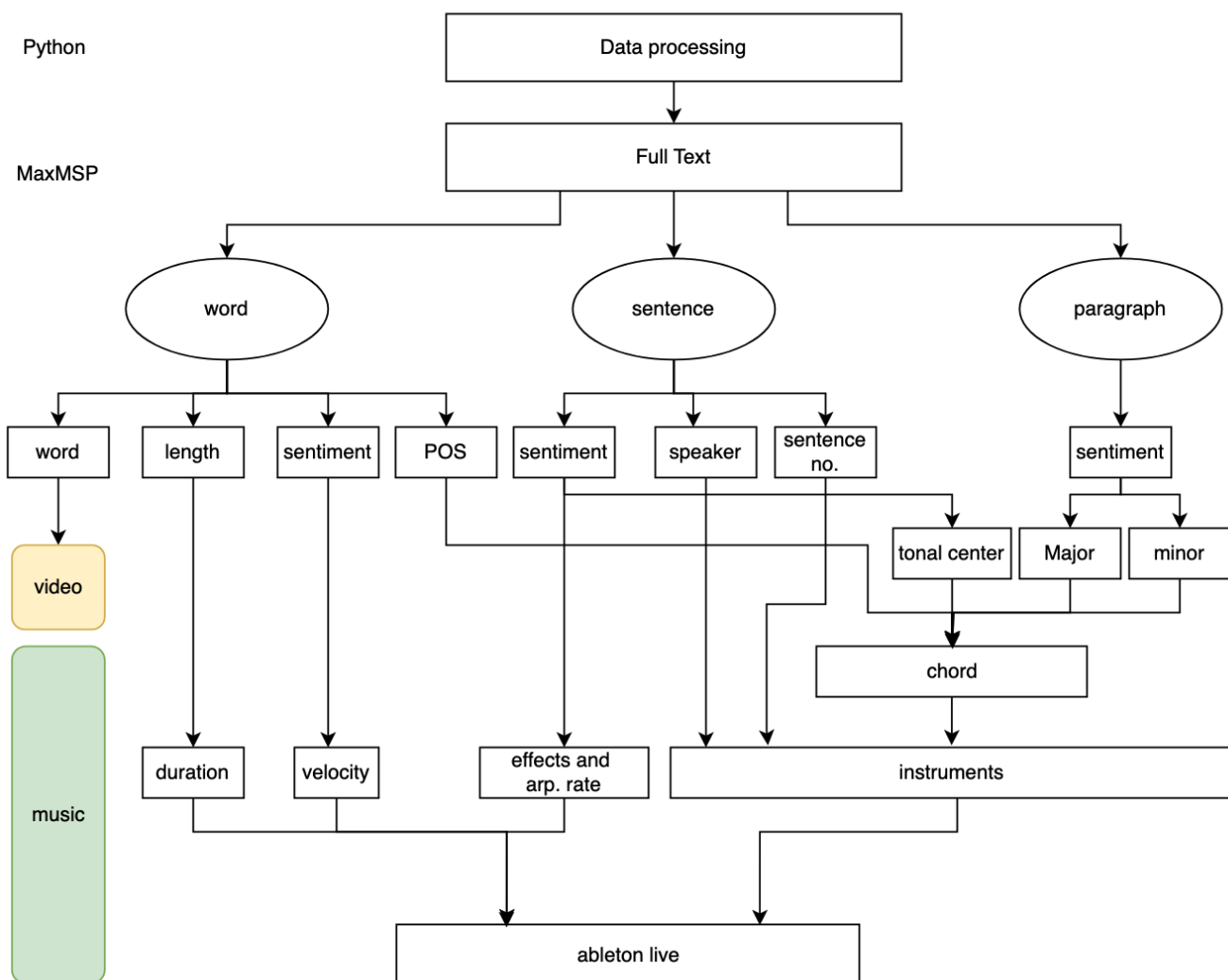


Figure 2: Algorithmic hierarchy of textual sonification

Sentiment information analyzed by Vader [13] controls musical parameters. Its scores shown from -1 to 1 in negative, neutral, positive and compound determine the emotional value of the text. **Word sentiment** index controls velocity of a note. **Sentence sentiment** (a) determines the tonal center where median value of the four indices is scaled and rounded up or down to decide the note to play from C to B in chromatic scale, (b) tweaks the arpeggiator rate, (c) controls attack time in envelope, and (d) controls delay amount where it is increased if the score is nearer to negative. **Paragraph sentiment** decides the key. Positive value from 0 to 1 goes to Major scale and negative value from -0.9 to -1 goes to minor scale.

5. DISCUSSION AND CONCLUSION

To maximize the effect of changing chords triggered by each word, the text is displayed digitally at intervals on a screen along with chords. The substituted chords generate unpredictable and riveting polyphony resulting in a composition that combines the inevitability and fortuity of language and music. A video of the excerpt of the piece titled *Animal Farm Sings* can be found online [1]. For a compar-

ative analysis, we visualized a sentence-POS plot for different sentences to see how mapped chord progressions can differ. Consider the well-known line from a poem, “Rose is a rose is a rose is a rose” [20]. Supposing the first Rose is a name of a person, it gives us the chord progression $iii - V - iib7 - I - V - iib7 - I - V - iib7 - I$. And a line from “Beasts of England” in *Animal Farm* which is written in the form of poetry gives us $I - ii - I - I - ii - I - I - ii - Iib7 - I - vii^{\circ} - I$. These simple chord progressions drawn from poems (Figure 3) are fairly different from that of general proses. To make the music more dynamic and articulate, POS can be subdivided into more detailed category. Verbs can be further categorized into regular verbs, linking verbs, auxiliary verb. Determiners and articles into possessive, indefinite / definite article and so forth. Other word types such as gerund, possessive, interjection can also be used as parameters. The attempt to link language and music has led us to find a way to use part of speech as chord progression that can turn text into harmonic music. The final composition bears the quality of contemporary music and is musical enough to be an independent piece of music. An approximately three-hour-long piece of music out of the text of *Animal Farm* was composed using automated algorithm. This textual sonification

Sentence	Chord	Key
1. Whatever goes upon two legs is an enemy.	iib7 - V - ii° - i - V - iib7 - i	-0.5423 (minor)
2. Whatever goes upon four legs, or has wings, is a friend.	Iib7 - V - ii° - I - vii° - V - I - V - Iib7 - I	0.4939 (Major)
3. No animal shall wear clothes.	iib7 - i - V - V - i	-0.296 (minor)
4. No animal shall sleep in a bed.	iib7 - i - V - V - ii° - iib7 - i	-0.296 (minor)
5. No animal shall drink alcohol.	iib7 - i - V - V - i	-0.296 (minor)
6. No animal shall kill any other animal.	iib7 - i - V - V - iib7 - VI - i	-0.7845 (minor)
7. All animals are equal.	Iib7 - I - V - vi	0.0 (Major)

Table 2: Chords and keys applied in The Seven Commandments

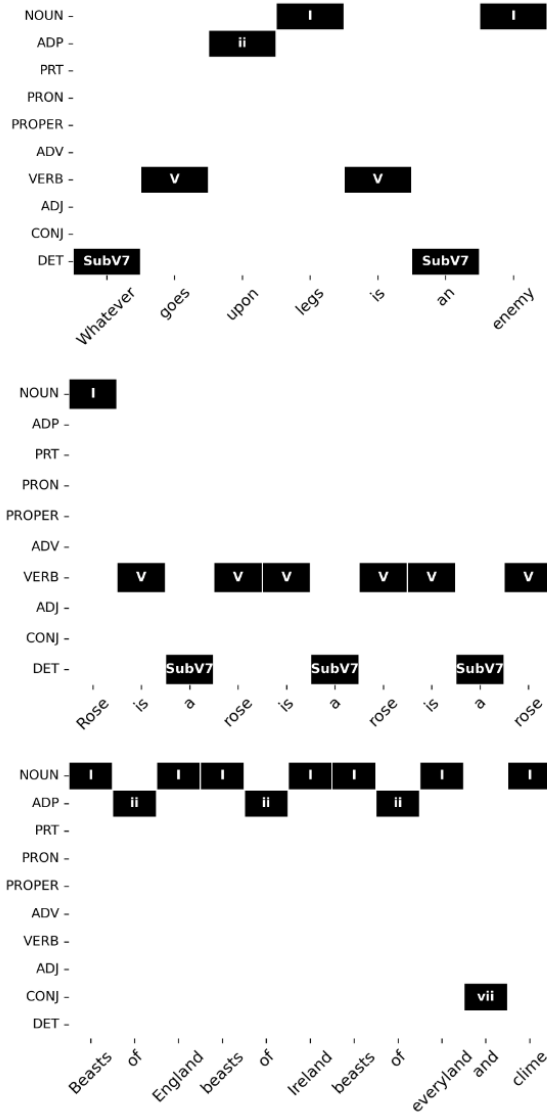


Figure 3: Comparison through graphical scores: (Top) A prose in *Animal Farm*; (Middle) A line from a poem *Sacred Emily*; (Bottom) A poetic line from *Animal Farm*.

enables us to grasp the context and harmony of the text with our eyes as well as our ears.

6. ACKNOWLEDGEMENT

A deep thanks to PyoungRyang Ko for his technical supervision and valuable support. This project was initially funded by the Arts Council Korea as part of Art and Technology Convergence Support Grant Program.

7. REFERENCES

- [1] Animal farm sings.
- [2] Melobytes.
- [3] A. Agostinelli, T. I. Denk, Z. Borsos, J. Engel, M. Verzetti, A. Caillon, Q. Huang, A. Jansen, A. Roberts, M. Tagliasacchi, et al. Musiclm: Generating music from text. *arXiv preprint arXiv:2301.11325*, 2023.
- [4] T. Almutiri and F. Nadeem. Markov models applications in natural language processing: a survey. *Int. J. Inf. Technol. Comput. Sci*, 2:1–16, 2022.
- [5] F. Alt, A. S. Shirazi, S. Legien, A. Schmidt, and J. Mennenöh. Creating meaningful melodies from text messages. In *NIME*, pages 63–68, 2010.
- [6] C. Barlow. Algorithmic composition, illustrated by my own work: A review of the period 1971–2008. In *Proceedings of Korean Electro-Acoustic Music Society’s 2011 Annual Conference (Korean Electro-Acoustic Music Society, Seoul, 2011)*, 2011.
- [7] S. Bird, E. Klein, and E. Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. ” O’Reilly Media, Inc.”, 2009.
- [8] T. Eagleton. *Literary theory: An introduction*. John Wiley & Sons, 2011.
- [9] J. Eastwood. *Oxford Guide to English Grammar*. Oxford University Press, Oxford, 1994.
- [10] E. Fedorenko, A. Patel, D. Casasanto, J. Winawer, and E. Gibson. Structural integration in language and music: Evidence for a shared system. *Memory & cognition*, 37:1–9, 2009.
- [11] Y. Hua. *Melody of Literature, Description of Music*. Prunsoop, Paju-si, 2019.
- [12] C.-F. Huang, H.-P. Lu, and J. Ren. Algorithmic approach to sonification of classical chinese poetry. *Multimedia Tools and Applications*, 61:489–518, 2012.
- [13] C. Hutto and E. Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the international AAAI conference on web and social media*, volume 8, pages 216–225, 2014.
- [14] R. Jackendoff. Parallels and nonparallels between language and music. *Music perception*, 26(3):195–204, 2009.
- [15] F. Lerdahl and R. S. Jackendoff. *A Generative Theory of Tonal Music, reissue, with a new preface*. MIT press, 1996.
- [16] E. Morales, K. James, R. Horst, Y. Takeda, and E. Yung. The sound of our words: Singling, a textual sonification software.
- [17] G. Orwell. *Animal farm*. Oxford University Press, 2021.
- [18] A. D. Patel. *Music, language, and the brain*. Oxford university press, 2010.
- [19] S. Petrov, D. Das, and R. McDonald. A universal

part-of-speech tagset. *arXiv preprint arXiv:1104.2086*, 2011.

- [20] G. Stein. Sacred emily. *Geography and plays*, pages 178–188, 1922.
- [21] E. Suhrob and K. Vasila. Parts of speech and sentence structure in english grammar. *Galaxy International Interdisciplinary Research Journal*, 10(7):156–160, 2022.
- [22] D. Tymoczko. Local harmonic grammar in western classical music. *Unpublished manuscript*, 2010.
- [23] A. Voutilainen. *Part-of-speech tagging*, volume 219. The Oxford handbook of computational linguistics, 2003.