Music, Notation and the Representation of Lexical Tone

Murray Schellenberg
University of British Columbia, Canada
mhschellenberg@gmail.com

Abstract: Music has played a significant role in the study and analysis of lexical tone. This paper presents a brief summary of some of the ways scholars have used and adapted conventions of Western musical notation to further their study of tone. It also examines the possible influence of cultural perceptions of melody on Western perspectives of tone.

1 Introduction
Western scholars have been aware of tone languages at least since the early 1600s [12]. Since that time, missionaries and scientists have devised a wide range of systems to describe and express these tones on paper. Music has played a central role in the way that Western scholars have described and understood lexical tone. Tone languages have long been associated with music in the West, with lexical tones frequently been equated with music in academic discussions of tone languages. For example, Crowther [8, p.3] describes Yoruba as “very musical”; Douglas [10, p. 33] says that Chinese tones “partake of the nature of musical intonations”; Werner [24] and Haupt [14] both use the term “musical accent” as an equivalent for “tone”; and Davis [9, p. 154] describes Navaho as “a singing language”. Cognitive scientists have also linked linguistic tones and musical pitch in cognitive processing (e.g. [16], [11]).

Early European scholars who wished to express lexical tones in writing faced a challenge. Western European orthographies represent speech sounds but they don’t represent the changes in pitch that characterize the speech act. However, when early scholars wanted to express pitch in language they had at hand a system that was designed expressly to represent both pitch and rhythm: musical notation.

This paper is not intended as an in depth discussion of all the various ways that tones have been described or annotated. Rather, the aim is to lay out a brief history of the diversity and potential usefulness of music-based tone/pitch annotation in speech. This paper will explore some of the ways in which the Western view of and approach to music have had an influence on the ways in which scholars from that same society have portrayed pitch in language. While the primary focus will be on the representation of linguistic tone, many of the same influences may also be found in other areas such as intonation. It should also be noted that both music and tone also involve duration as well as pitch and that representations of duration can be found in many of the descriptions described in this paper. For reasons of brevity, however, the discussion in this paper will be limited solely to pitch.

Section 2 will present a very brief resumé of the development of musical notation in Western Europe; sometimes referred to as the first technology for capturing pitch in the West [15]. The subsequent sections will introduce different ways that scholars have used Western musical notation to represent lexical tones; starting with attempts to transcribe speech tones directly into musical notation and moving through various adaptations that have been used. The final section gives a brief discussion of the idea of melody as a cultural construction and offers some suggestions on how the prevailing Western European conception of melody may have had an influence on the academic interpretation of lexical tones.
2 Western Musical Notation – A system for capturing sound

Sound is ephemeral – once it is made, it disappears completely. For thousands of years the only way to reproduce a sound was to remember it and then to make it again. Writing was the first technology for recording sound with written language long predating written music in the West [15, 21]. The earliest songbooks in medieval Europe were just books of words used for the many religious services in monasteries. Monks spent a good portion of their day chanting – and the chants were often different from one day to the next. Without a system for notating the melodies, the only way for the monks to learn the melody for a particular chant was to hear it and then try to remember it [15]. While some chants were repeated fairly frequently, others were sung only once a year or even less often. At some point, a monk with a poor memory (probably various monks in different places) started to write in reminders of the melody above the words in the songbook. Soon these reminders were being copied into new books that were being produced. Figure 1 shows an example of this kind of notation from a breviary (book of liturgical texts) produced at the Abbey of St. Gall in Switzerland between 1022 and 1047. This early notation consists of lines, dots and squiggles (collectively called neumes) that give indications of how the melody moves. As the words were already written from left to right, the music notation acquired the same left-to-right alignment. Pitch level within a neume was indicated via higher pitches being closer to the top of the page. Sometimes the relative pitch level of adjacent neumes was indicated by vertical placement on the page … but often it wasn’t [15].

The neumes, it is important to note, served only as reminders of a melody already known. Someone who didn’t know the melody would not be able to reconstruct what it should sound like by reading the neumes. If you wanted to learn it, you still had to find someone who already knew the melody to teach it to you.

That changed thanks to a man called Guido d’Arezzo, a Tuscan monk who lived from around 995 to about 1050 [18]. Somewhere around 1030 Guido devised – or at least, wrote down – a new system for recording music. This system involved drawing a grid of 4 horizontal lines and superimposing the neumes on it. Grid lines had been used prior to this [12] but Guido’s real innovation was to specify that any neume written on a particular line or space of the grid would always have the same (relative) pitch. He also specified the relationships between the pitches – the distance between adjacent notes is not always the same [15]. His pitch relationships are still used today as our modern day major scale and his gridlines were the forerunner to the modern musical staff. This innovation led to a major change in the usability of musical notation: it now became possible for someone who had never heard a song before to learn it without a teacher [15]. An example of this type of notation from a 15th century Italian gradual (book with the music for the mass) is shown in figure 2. Over time, musicians
have made further refinements to the system but the basic 5-line staff and pitch relationships of notes used by musicians today have their direct antecedents in the innovations of Guido d’Arezzo.

3 Using Musical Staff Notation for Tone

Remarkably few scholars have used the Western musical notation system directly to represent lexical tones but it has been done. This section presents two examples: one from a description of Thai in the mid-nineteenth century and one from a Mandarin grammar from the 1960s.

The first example is the 1850 grammar of Thai [20] written (in Latin) by Denis Jean Baptiste Pallegoix who was the Bishop of Siam (as it was then called). The chapter on tone is brief (8 pages in a book of 241 pages) but in it, he provides typographic representations of the tones via accent markings. He also (p. 36) provides musical transcriptions of not only the tones in isolation (as shown in figure 3) but also provides (loc. cit.) musical transcriptions of very brief samples of colloquial speech (modulus in colloquio communi), prose recitation (in recitation prosœ), preaching (in prœdicatione) and singing (in recitatione carminum). Pallegoix’s transcriptions would have to have been made based on his own perceptions of native speakers. They are somewhat different from the modern-day representation of the Thai tone contours (c.f. figure 9, for example) but certainly express distinct differences. They would also have been easily perceived and produced by anyone able to read music.

The second example comes from Yuen Ren Chao in his 1965 Grammar of Spoken Chinese [7]. In addition to written descriptions of the tones and their representation in the IPA tone letters (see 5.1 below), Chao provides a brief representation of the tones transcribed musically (Figure 4). Western descriptions of Mandarin tones have assigned numbers to the four tones and Chao presents the tones in that numerical order. Unlike Pallegoix, Chao does not provide sample words beneath the tones but presents the musical score independent of the words.

Figure 2. Music with fixed relative pitches on a 4-line staff. Excerpt from Alleluia for mass on Christmas Day – 15th century Italian manuscript. UBC, Italian 15th century gradual [M]. http://dx.doi.org/10.14288/1.0054703

Figure 3. Musical notation used to explain tone in Thai. The typographic tone marks can be seen on the vowels of the words under the score [20, p. 36]
It is interesting that Chao – who is as renowned a composer as he is a linguist – chose to add musical notation at a time when other graphic representations such as pitch contour graphs were also readily available. This choice probably reflects what is one of the major advantages of using musical notation: it is a very widely understood system. It is immediately accessible to a wide range of people and has such a common base of understanding that it takes little to no explanation for a very large proportion of people to get a clear picture of both the pitch contours and relative pitch levels.

Music notation, however, is not necessarily very accurate for the representation of speech. Pitch and melody in (Western) music are not really the same as pitch and melody in speech. In 1779, Steele [22] noted that:

[Music is] a series of sounds moving distinctly from grave to acute, or vice versa (either gradually or saltim) by intervals, of which the semi tone (commonly so called) may be the common measure or divisor, without a fraction, and always dwelling, for a perceptible space of time, on one certain tone.

Whereas the melody of speech moves rapidly up or down by slides, wherein no graduated distinction of tones or semitones can be measured by the ear; nor does the voice (in our language) ever dwell distinctly, for any perceptible space of time, on any certain level or uniform tone, except the last tone on which the speaker ends or makes a pause. (p. 4)

Steele [22] was just one of the scholars who decided to take standard musical notation and modify it to make it more useful for depicting pitch in speech.

4 Modifications of Staff Notation
In his 1779 book, *Prosodia Rationalis: Or, An Essay Towards Establishing The Melody And Measure Of Speech, To Be Expressed And Perpetuated By Peculiar Symbols*, Joshua Steele lays out a system to notate the prosody of speech so writers and orators could notate the prosodic component of speech. While his system was not intended to be used for tone languages it is worth noting in the context of tonal representations as he takes the syllable as the basic unit for pitch notation and notates contours for each syllable.

Significantly, Steele explicitly chose to make his system a modification of the Western musical notation system. He explains [22, p. 8] that he decided to develop a system “which might come as near as possible to the modern notation of music, in order to make it the easier to be comprehended by those whose ideas of sounds and measure of time are already formed on that plan.” But he needed to find a way to express the “slides” of speech which exhibited “no gradual distinction of tones” [22, p.4].

An example of his notation is given in figure 5. His system is based on the 5-line staff of modern music notation, but each space between the lines is further subdivided to provide a line for each semitone in the Western scale. The pitch contour of each syllable is drawn as a dark line and the various tails attached to each contour indicate different rhythms (the calli-
Over a century later, in the 1930s, Douglas Beach devised a very similar system (described in [23]) of expanding the standard 5-line staff to produce a massive 29-line staff to encompass a pitch range of an octave and a third (figure 6b). It is interesting to note that his subdivisions of the original 5-line staff precisely matches that of Steele with the outer spaces of the staff being subdivided into four and the inner spaces being subdivided into three (see figure 6) Beach added a further 7 lines above and 7 lines below. Tucker [23, p.596], who was a student of Beach, relates that “students, armed with tuning-forks, were set to record on these sheets accurate representations of tones as heard from phonograph cylinders, picking the needle up at each syllable.”

These expanded staves provide a somewhat finer representation of pitch than the conventional staff notation of music. As each line represents a semitone, it is possible to notate pitch down to quarter tones by using the space between the lines to indicate the space between the notes [22, p.5]. The density of the lines, however, makes rapid reading much more difficult. Both Steele and Beach tried to alleviate this somewhat by changing the density and pattern of the lines but it is still difficult to read quickly. It is possible, for example, to determine that the rise on the first syllable in Figure 5 is 15 semitones – but only after careful counting. Determining accurate pitch changes quickly is tricky.

---

**Figure 5.** Example of Steele’s proposed speech prosody notation [22, p. 13]

---

**Figure 6.** Comparison of expanded staves used by (a) Steele [22, p. 5] and (b) Beach [23, p. 596]
5 Staffless Pictorial Representations

As absolute pitch is not generally necessary for depictions of tone, pictorial representations of tone have often abandoned the musical staff. The traditions of Western music, however, have had various influences on these representations, as well. For example, all of the systems discussed below (and many others, as well) mirror two of the most basic conventions of musical notation: time is portrayed along the horizontal axis from left to right and pitch is portrayed along the vertical axis with higher pitches being shown as closer to the top of the page. As well, scholars have regularly employed musical note names and musical intervals to express pitch levels and ranges.

5.1 Diagrammatic systems accompanying orthography

The tone letters used as part of the International Phonetic Alphabet were first suggested in an article by Yuen Ren Chao in 1930 [6]. The letters consist of a vertical line divided into 4 equal parts, thus providing 5 numbered points. A “time-pitch curve” is drawn to the left of the line indicating the shape and pitch level(s) of the syllable’s tone (e.g. ˧˥ , ˨˧˦). The musical influences are not, perhaps, as readily apparent from the symbols but the time-pitch curves very definitely mimic the time/pitch relationship of musical notation. Furthermore, the divisions on the vertical line are based on musical intervals. Chao [6, p. 25] stipulates that “each step may be taken to be a whole tone, thus making the total equal to an augmented fifth.” The divisions of tone steps in this system are equated directly to musical intervals.

Another pictorial system uses dashes and lines to express tone movement in language (see Figure 7). Ida Ward made extensive use of one such system in a series of African language primers in the 1930s [23]. This system involved using vertically positioned dashes to indicate pitch height and curves to indicate rises or falls.

Tucker [23, p. 600] describes it as “an eminently pedagogic system, and not intended for indigenous authors; but she [Ward] felt that the Westermann system [an orthographic system of accent marks] lacked sufficient tone marks for the necessary levels of down step which she had to teach. Her system’s main disadvantage was its expense, though the English printers were very accommodating and allowed her five levels of horizontal line per block.” The resulting tone descriptions look remarkably like the neumes in the hymn books of the medieval monks (Figure 1).

Figure 7. Ward’s diagrammatic system for 3 words in Twi (as shown in [23])

Figure 8. Caswell’s 1870 diagram of Thai tones. [5, p. 95]. He describes Thai as having only 4 tones (C, D, E and F). The line AB is the “pitch of the voice in reading English”
5.2 Pitch contours

The tradition of horizontal time and vertical pitch are also found in the linear pitch contours, a representation frequently used to portray tones today. Figure 8 shows an early diagram of this sort from 1870 [5]. This is another description of the tones of Thai. Caswell [5] describes Thai as having only 4 tones (conflating the two level tones). His description runs “Let the line AB represent the natural key, or pitch of the voice in reading English … C will represent the rising inflection, D the falling inflection, E the depressed tone, and F the circumflex tone.” [5, p. 95]. He uses the idea of a reference line not dissimilar to a musical staff line around which he can arrange pictorial descriptions of the four tones.

The influence of music continued with the advent of mechanical means of recording pitch. Cornelius Bradley was the first to use the kymograph to record the production of tones [4]; first in Thai [2] and then in Cantonese and Mandarin [3]. In 1916 he published a monograph explaining the procedures he used to produce his pitch contours [4]. The kymograph was used to record the waveforms of the tones; fundamental frequency values were then computed from hand measurements taken off the waveform and subsequently graphed. Figure 9 shows his 1911 graph of the tones for Thai [2]. Note that the pitches are given as musical note equivalents along the vertical axis despite the measurements all having been made in cycles per second. Musical notes are still an easy and convenient way of making pitch values accessible.

![Chart of the Five Tones of Long Syllables in Siamese.](image)

**Figure 9.** Bradley’s pitch contours of Thai Tones based on kymographic recordings [2]
Pitch and the Conception of Melody

The preceding examples have all shown ways in which the system of musical notation prevalent in the West has influenced, to varying degrees, visual representations of pitch (primarily tone) in language. This section takes a somewhat more ethnomusicological approach and attempts to look at what influence the culturally specific Western-European conception of melody has had on the perception of tone by scholars from that culture.

Any musical system is a cultural construct which has modified and molded the auditory signal in ways that reflect what that culture deems to be important. Nettl [19, p. 85] points out that “there may be, in certain cultures, segments of sound more significant than those we label as ‘notes’”. Ethnomusicologists have wrestled for years with the difficulty of defining music and describing, for a primarily Western audience, non-Western conceptions of music [19]. While a given acoustic signal may be non-cultural, its perception by an individual is highly coloured by cultural factors [19, 21] – which have been described as “norms determined by the vast aggregate of practice and codified by generations of workers” [21, p. 193]. While Seeger in that description is referring specifically to notation, his description applies equally well to the development of any cultural norm.

The conception of melody is an excellent example of something that is open to cultural interpretation. Seeger [21] proposes that melody is something that may be conceived in a variety of ways. He suggests that “on one hand ... melody may be conceived ... as a succession of separate sounds, on the other, as a single continuum of sound” (p. 185). He further divides the conception of melody as a continuum into two different possibilities: “as a chain or as a stream. Conception as a chain tends to emphasize structure and entities that move; conception as a stream, function and movement” (p. 185).

The concept of melody in the West is, of course, very much a conception of discrete units in a sequence. This can be seen in the very idea of musical “notes” so prevalent in the music of Europe and North America. The early development of neumes into representations of individuated pitches suggests that Western cultures have long conceived melody as consisting of discrete pitch units.

The Western conception of melody as a series of individual pitch targets has interesting implications for our understanding of how we perceive pitch and melody in speech. Lexical tones, for example, are theoretically conceived in Western linguistics as discrete pitch units: high and low. It is possible to see the idea of lexical tones as pitch targets as being influenced (perhaps unintentionally) by the prevailing musical conceptions of melody and pitch. Did we simply transfer over our musical perceptions to language?

While this perception/construction of tone has allowed the development of extremely useful insights into language it is not the only way to conceive of tone. Chinese scholars at the time of Middle Chinese (roughly 420 – 1273 C.E.) categorized the tones of Middle Chinese as level, rising, departing and entering [1, 16]. Mei [17, p.91ff] provides a fascinating translation of a 7th century Japanese description of these four Chinese tones:

The level tone was level and low, with both the light and the heavy; the rising tone was level and high, with only the light but not the heavy; the departing tone was slightly drawn out, with no [distinction between] the light and the heavy; the entering tone stops abruptly, having neither the inner nor the outer; the level tone [carried by syllables] with nasal or lateral initials was indistinguishable from the heavy; and the heavy of the rising tone was no different from the departing tone. (annotations from [17])
While an in-depth discussion of the Middle Chinese tone system is well beyond the scope of this paper, it is interesting to note that the names of the tones (albeit in translation) appear to have an emphasis on movement and contour and not on pitch height. Scholars interested in tone in Middle Chinese appear to have been more influenced by poetry than music [1, 16] and their approach to categorizing tones is quite different to that of modern Western scholars.

7 Conclusion

Music has had a continuing influence on how Western scholars represent lexical tones. Although musical notation itself acquired some of its conventions from language (primarily the left-to-right use of the page), conventions of musical notation such as pitch height being towards the top of the page have become standard in pictorial depictions of lexical tone. The widespread use of musical notation culturally has meant that use of this notation has made representations of lexical tone easily accessible and understandable to a population that does not speak a tone language. However, the differences between speech and music have meant that adaptations have often been desirable. Scholars have adapted musical notation in various ways but its conventions have often prevailed. Even when tone notation systems have appeared to be quite different from music notation, scholars have often relied on musical note names and pitch intervals to make their work more readily accessible and understandable.

As well, the Western European construction of melody as a series of individual notes has possibly also been influential in the prevailing idea of lexical tone being constructed of individual units. As mentioned before, this has led to many fascinating insights but it is possible that additional insights might be made by reconceiving tones in a “movement” conception.

This paper has tried to show that music has had a regular and important influence on Western linguistic research into lexical tones. Musical notation has provided a ready means by which scholars have been able to record and preserve sound and make it available to others. Often this influence has gone unrecognized but awareness of these connections may allow future scholars to bring new insights to the field.

8 References

[22] Steele, J.: Prosodia rationalis: or, an essay towards establishing the melody and measure of speech, to be expressed and perpetuated by peculiar symbols. London: J. Nichols, 1779