Voice-leading in Palestrina's masses: A comparison of interval-succession definitions

Laney Light[1†]
Claire Arthur[2]
[1] School of Music, Georgia Institute of Technology, Atlanta, GA, USA
[2] School of Music, Georgia Institute of Technology, Atlanta, GA, USA
[†] Corresponding author: llight7@gatech.edu

Abstract

Computational musicology methods allow us to perform a systematic analysis of all intervals in a corpus, but are all note successions equally important? In traditional music theory, the pulse beats tend to be given greater "weight" or relevance. In Renaissance music, it is unclear whether voice-leading guidelines should apply to any consecutive intervals (note level), or merely those that traverse from one pulse beat to the next (pulse level, defined by the whole note or semibreve). Since this question bears important implications for computational musicology, we set out to empirically evaluate it via a corpus study of Palestrina's masses.

We investigated the voice-leading patterns in Palestrina's corpus of 104 masses using music21. For each pair of voices, we systematically investigated whether Palestrina's voice-leading patterns differed at the note level compared to the pulse level. Our results showed that the distribution of voice-leading patterns was significantly different at the pulse level than the note level. Violations of traditional "rules" prohibiting parallel or similar motion between perfect intervals were more common at the pulse level than the note level ($p < .05$), while factors associated with breaking the rules (e.g., having more voices in the texture) were similar between the two levels.

KEYWORDS: *Voice leading, Renaissance, Palestrina, corpus study*

Introduction

Voice-leading, an integral aspect of Western music, refers to the way individual voices, or musical parts, move together in a multi-part texture. This movement must consider each voice's melodic motion (horizontal intervals) as well as the harmony between voices active at the same time (vertical intervals). Voice leading in Western music originated in early forms of polyphonic singing in the Middle Ages and evolved into complex counterpoint in the late Renaissance (Fuller, 2002). Treatises by Renaissance music theorists included guidelines for what was generally considered "good" and "bad" polyphonic writing. However, there were seldom explicit "rules" as one might find in a modern theory textbook (e.g., Aldwell et al, 2019) or a 20th century text on writing in the Renaissance style (e.g., Gauldin, 1985). Guidelines in these modern textbooks are developed from the authors' internalization of "inherent rules" abstracted from years of study of early musical treatises and familiarization with a vast body of early contrapuntal composition. However, given the magnitude of musical material from this period, it is all but impossible to examine voice-leading practices in a systematic manner using traditional humanistic methods alone. In this paper we comprehensively examine the voice-leading practices in a large body of Renaissance polyphony to clarify certain ambiguities (or inconsistencies) in ancient and modern texts.

Examining voice-leading patterns in Renaissance music using computational methods has implications for Renaissance voice-leading pedagogy because it can highlight subtle variations between Renaissance theory and practice, clarify the context under which certain rules operate, and potentially identify a set of more general principles guiding certain compositional behaviors. Since Renaissance polyphony formed the basis for voice-leading practices through the Baroque, Classical, and into the modern era, any insights gained from this research are likely to have implications for voice leading in later musical styles as well.

Although modern computational methods provide powerful tools for analyzing scores, defining compositional rules algorithmically requires careful consideration. For example, it is often unclear how the "rules" of voice leading should be applied regarding meter. Isolated musical examples—especially in Renaissance treatises—are often presented in a prototypical notation that appears agnostic to the "beat" (or position), duration, or voice position, or else are demonstrated with a very specific selection of musical material. One question that arises, then, when trying to interpret these examples is whether they represent a rule to be followed at any "hierarchical level?" Or, for example, whether there are separate rules that would apply at the "note level" (any consecutive intervals)

versus the "pulse level" (from one strong beat to the next). That is, are voice-leading rules to be applied differently at different hierarchical levels?

Andrews (1958) describes a voice-leading "violation" where direct consecutive octaves at the pulse level are (ineffectively) interrupted by a passing tone. However, we do find evidence of this pattern (and other "violations") used occasionally in Palestrina's masses, as illustrated in Figure 1. Using computational techniques, we can analyze millions of individual intervals to quantify how often these types of patterns occur in compositional practice, and under what circumstances.

Figure 1. Excerpt from Palestrina's Missa Sine nominee (Mantuan), Agnus. The prohibition on direct consecutive perfect intervals is violated at the pulse level (*) but not at the note level because of the passing tone (+).

In this study we take a computational approach to investigate whether Renaissance compositions differ in their voice-leading treatment at the note level versus the pulse level. To answer this question, we used a digitally encoded corpus of Palestrina's masses. This corpus is ideal for this analysis since Palestrina is well known as an exemplar of Renaissance vocal polyphony (Benjamin, 2005), and therefore his work should be representative of the voice-leading practices of the time. Moreover, given Palestrina's huge influence on the development of contrapuntal practice (Marvin, 2002), we would expect to find evidence of similar voice-leading practices in later polyphonic music.

Related Works

There have been many studies on Renaissance counterpoint and Palestrina in particular (e.g., Jeppesen,

1927; Marshall, 1963; Hanson, 1983). These have covered a wide range of historical and music theory topics, but most have used traditional musicology approaches. Using computational methods, we can build on these theoretical works to examine scores systematically and identify patterns from large datasets in ways that are not possible with traditional approaches.

Palestrina's compositions provide a rich data source for computational analysis because of the large number of works that have been encoded as symbolic score data. Several studies have leveraged this data source for computational analyses (e.g., Arthur, 2021; Sigler, 2015; Knopke et al, 2009; Farbood & Schoner, 2001), but to date only Arthur (2021) has examined Palestrina's voice-leading specifically.

There have been a few computational studies focusing on voice leading. Wall et al (2020) performed an empirical study of the effects of voice leading and harmony on musical expectancy, but the perceptual study did not rely on data from symbolic scores. Huron and Collins (1999), whose work most resembles our own, investigated the degree to which voice leading guidelines set by Zarlino and Berardi agreed with compositional practice. As it was unclear whether melodic rules should apply to all intervals or only those on the strong beats, the authors created two separate "inventories" of melodic intervals, one containing note-to-note intervals and the other containing intervals between notes on the strong beats. However, their work only examined voice-leading in a stricter and highly imitative style (canon for 3 voices). In addition, while their sample was formidable in size and scope (79 canons and 13 didactic examples by 13 composers over a period of 250 years), the overall sample of intervals is comparatively small due to the short length of the canon form itself. Most recently, Arthur (2021) performed a comprehensive examination of Palestrina's voice leading to examine the role of vocal texture on voice-leading rules and preferences.

How intervals are represented (or tallied) affects one's results. There are many ways in which voice-leading patterns can be represented symbolically. For example, Conklin (2002) fully expanded each piece, creating a new vertical "slice" at each onset to produce a "viewpoint representation" that can be sampled at regular intervals (quarter notes, for example). Sears et al (2002) presented a "skip-gram" representation. Based on the n-gram representation (e.g., a trigram is a sequence of $n = 3$ events), the skip-gram is defined as non-adjacent sequences that skip over n number of events. This method, which was extended by Finkensiep

eta al (2018) to allow for nested skip-grams, can uncover higher-level structural patterns that may be "hidden" by other passing tones. However, the skips are not dependent on the beat positions of the notes. Arthur (2021) only examines voice-leading as a function of simple 2-gram (i.e., note level) patterns. In our analysis, we build on Arthur's work, but examine several of these representation models (vertical slices from a full expansion of the score, sampled at the note (bigram) level and at the pulse beats to obtain two separate inventories of intervals).

Method

Voice-Leading Patterns

We used a systematic computational approach to analyze voice-leading patterns in a corpus of 104 masses by Palestrina using music21 (Cuthbert, 2010) [1]. We performed a full expansion of the scores, "slicing" the scores vertically at any new onset in any voice. Next, we calculated the vertical bigrams (harmonic intervals) between every pair of voices in each slice, as well as the horizontal bigrams (melodic intervals) for each voice. Then we categorized the types of contrapuntal motion leading to each vertical interval, as shown in Figure 2: parallel, contrary, similar, oblique, and stasis (one or both voices resting on the previous vertical interval, or both voices repeating the same notes).
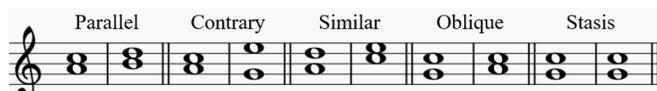

Figure 2. Examples of voice-leading motion types.

We created two non-mutually exclusive inventories of intervals, "note level" and "pulse level" [2]. Within each inventory, we identified a set of features associated with each interval, including the position of each voice within the texture (inner vs. outer voice), total number of voices in the texture at that slice, and whether each interval was approached by a step or a leap. Next, building on the work of Arthur (2021), we searched each inventory for specific patterns that violate two voice-leading "rules" governing harmonic intervals.
• R1: perfect harmonic interval (P5, P8, or P1) should not be approached by parallel motion
• R2: perfect harmonic interval (P5, P8, or P1) should not be approached by similar motion

Statistical Analysis

We ran statistical tests to assess the associations between metric position, hierarchical levels, and voice-leading patterns. All statistical tests were conducted using JMP® software [3], with an alpha level of .05.

First, we tested whether the usage of certain types of harmonic intervals (consonances and dissonances) differed for pulse-beat onsets compared to other metric positions. We categorized the harmonic intervals into perfect consonances (P1, P5, and P8), imperfect consonances (m3, M3, m6, and M6), and dissonances (all other intervals). Using a Pearson chi-square test, we compared the distributions of the interval categories for harmonic intervals landing on a pulse beat, compared to all other metric positions.

Next, we evaluated whether Palestrina used different voice-leading patterns at different hierarchical levels. To do this, we compared the distribution of motion types (parallel, contrary, similar, oblique, and stasis) in the note-level inventory vs. the pulse-level inventory, using a chi-square test. We also calculated the percentage of intervals that violate each of the above rules *within* each inventory, using a chi-square test to determine whether the prevalence of 'rule violations' differed at the pulse level and note level.

Finally, we investigated the specific conditions under which the rules were violated using logistic regression models, with separate models for each voice-leading rule and for the note-level and pulse-level inventories. The dependent variable for each model was whether the rule was violated (yes vs. no). The independent variables were: relative position of the consequent interval (inner/inner, outer/outer, or inner/outer); total number of voices in the texture at the consequent interval (categorized as 2 to 3, 4 to 5, or 6+ voices); whether the upper voice moved by a leap; and whether the lower voice moved by a leap.

Results

The note-level inventory contained 1,705,371 harmonic intervals with their associated features, which was approximately 2.8 times the size of the pulse-level inventory (615,730 intervals).

The distribution of perfect, imperfect, and dissonant intervals on the pulse beats differed significantly from the distribution at other metric positions, $X^2(2, N = 1,705,371) = 109,993$, $p < .05$. As shown in Figure 3, consonances were more common on the pulse beats compared to other metric positions, consistent with trends described by Andrews (1958, p. 63).
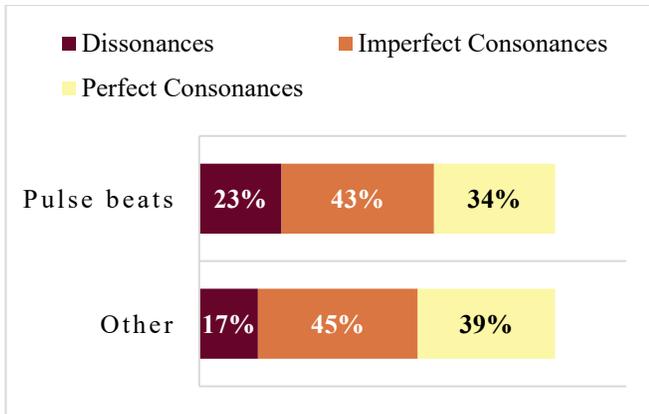
Figure 3. The distribution of harmonic interval types varies by metric position, with consonant intervals more prevalent on the pulse beats.

Voice-leading patterns also differed at the note level compared to the pulse level (see Figure 4). Similar, parallel, and contrary motion were more prevalent at the pulse level, while stasis and oblique motion were more prevalent at the note level, $X^2(4, N = 2,321,101) = 116,356, p < .05$.
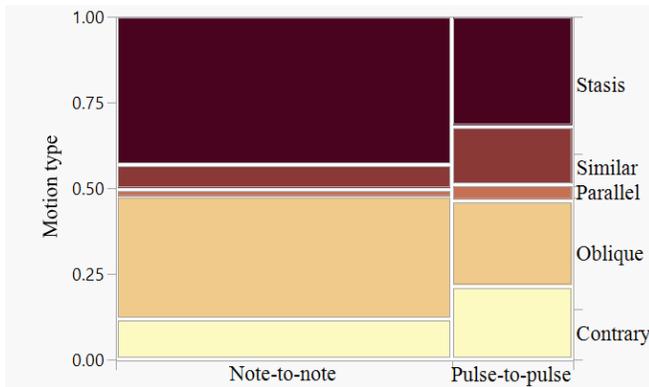

Figure 4. Similar, parallel, and contrary motion are more prevalent at the pulse level than the note level. The column width is proportional to the total number of intervals in the inventory.

Both voice-leading rules (R1 and R2) were broken significantly more often at the pulse level than the note level. We observed that R1 (parallel motion to perfect interval) was almost never broken at the note level (0.01%), but this voice-leading pattern was used 2.2% of the time at the pulse level, $X^2(1, N = 2,321,101) = 35,976, p < .05$. R2 was also broken more often at the pulse level than the note level (3.6% vs. 1.2%), $X^2(1, N = 2,321,101) = 13,387, p < .05$.

The factors associated with breaking both voice-leading rules were similar at the note and pulse levels, although the effect sizes tended to be smaller at the pulse level ($p<.05$ for all effects). Outer/outer voice pairings were more likely to break both rules than inner/inner pairings. At the pulse level, outer/inner pairings were *more* likely to break R1 at the pulse level but *less* likely at the note level, compared to inner/inner voice pairings. Having larger numbers of voices in the texture and moving by leaps (as opposed to steps) were associated with greater likelihood of breaking both voice-leading rules.

Table 1. Parameter estimates from logistic regression models. The dependent variable for each model is rule violation (yes vs. no).

	Rule 1		Rule 2	
Effect	**Note level**	**Pulse level**	**Note level**	**Pulse level**
Position				
outer/outer (vs. inner/inner)	1.00	0.47	0.63	0.42
outer/inner (vs. inner/inner)	-0.54	0.39	0.36	0.25
Number of voices				
6+ (vs. 2 to 3)	0.82	0.06	0.30	0.25
4 to 5 (vs. 2 to 3)	0.20	0.19	0.16	0.04
Lower voice movement				
leap (vs. step)	0.48	0.08	1.70	0.90
Upper voice movement				
leap (vs. step)	0.79	0.28	0.58	0.16

Discussion

We found that consonances were more prevalent on the pulse beats, consistent with Andrews' (1958) observation that intervals on pulse beats are usually consonant (p. 62). Our analysis of voice-leading patterns showed that contrary, similar, and parallel motion types made up a larger proportion of the total intervals at the pulse level than at the note level.

Specific voice-leading "rules" (approaching perfect intervals by parallel or similar motion) were broken *more* often at the pulse level. This is seemingly contrary to what we could expect based on theoretical texts. Assuming any rule violations would be more obvious at perceptually salient levels, we would have expected fewer violations at the pulse level, "hiding" them instead at other beat positions.

We also identified features that were associated with a greater likelihood of breaking certain voice-leading rules. Both rules were more likely to be broken in textures with larger numbers of voices, supporting Arthur (2016)'s observation that it becomes more difficult to follow the rules as the number of voices increases.

Conclusion

We found that voice-leading patterns in Palestrina's masses differ at the pulse level compared to the note level. Our findings are relevant for voice-leading pedagogy because they suggest that Palestrina used certain types of "forbidden" voice-leading patterns more often than previously assumed, albeit at a higher metric level. This also has implications for other types of computational analyses using note-to-note successions, because using higher-level hierarchical structures could uncover different patterns or relationships. However, selecting the best unit of analysis for voice leading is still unclear. The most perceptually salient beats should be used for the pulse level analysis, but determining which beats are the most salient depends on many factors, including the tempo at which a piece would have been performed, which is difficult to determine from the symbolic music alone.

End Notes

[1] The pieces are encoded in *kern* format and publicly available via music21's built-in corpus.

[2] We defined the pulse as the semibreve (or whole note). While the pulse level can change (Morgan, 2016) DeFord notes that the semibreve is the most common case (DeFord, 2015, p.375).

[3] JMP®, Version 15. SAS Institute Inc., Cary, NC, 1989-2020.

References

Aldwell, E., Schachter, C., & Cadwallader A. (2019). *Harmony & Voice Leading* (5th ed.). Cengage.

Andrews, H.K. (1958). *An Introduction to the Technique of Palestrina*. Novello and Company.

Arthur, C. (2021). Vicentino versus Palestrina: A computational investigation of voice leading across changing vocal densities. *Journal of New Music Research*. Advance online publication. https://doi.org/10.1080/09298215.2021.1877729

Benjamin, T. (2005). *The Craft of Modal Counterpoint* (2nd ed.). Routledge. https://doi.org/10.4324/9780203997208

Conklin, D. (2002). Representation and Discovery of Vertical Patterns in Music. In C. Anagnostopoulou, M. Ferrand, & A. Smaill (Eds.), *Music and Artificial Intelligence* (Vol. 2445, pp. 32–42). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45722-4_5

Cuthbert, M. S., & Ariza, C. (2010). Music21: A toolkit for computer-aided musicology and symbolic music data. *11th International Society for Music Information Retrieval Conference* (ISMIR), Utrecht, Netherlands, 637-642.

DeFord, R. (2015). *Tactus, Mensuration, and Rhythm in Renaissance Music*. Cambridge University Press. https://doi.org/10.1017/CBO9781107587717

Finkensiep, C., Neuwirth, M., & Rohrmeier, M. (2018). Generalized skipgrams for pattern discovery in polyphonic streams. *19th International Society for Music Information Retrieval Conference* (ISMIR), 547–553.

Fuller, S. (2002). Organum – Discantus – Contrapunctus in the Middle Ages. In T. Christensen (Ed.), *The Cambridge History of Western Music Theory* (pp. 477-502). Cambridge University Press. https://doi.org/10.1017/CHOL9780521623711.017

Huron, D., & Collins, D. (1999). Voice leading in cantus firmus-based composition: A comparison between theory and practice in Renaissance and Baroque music using computer-assisted inferential measures. *Computers in Music Research*, 6 (Spring), 53–95.

Gauldin, R. (1985). *A Practical Approach to Sixteenth-Century Counterpoint*. Prentice Hall, Inc.

Hanson, J. R. (1983). *Enumeration of Dissonance in the Masses of Palestrina*. College Music Symposium, 23 (1): 50–64.

Hanson, John. (1983). Enumeration of dissonance in the masses of Palestrina. *College Music Symposium*, 23 (1): 50–64. College Music Symposium, 23(1), 50–64.

Jeppesen, K. (1927). *The Style of Palestrina and the Dissonance*. Translated by Nargaret N. Hamerik. Oxford University Press; revised English edition, 1946; reprint of the 1946 edition: Dover Publications, 1970.

Marshall, R. (1963). The paraphrase technique of Palestrina in his masses based on hymns. *Journal of the American Musicological Society, 16*(3). https://doi.org/10.2307/829827

Marvin, C. (2002). *Giovanni Pierluigi Da Palestrina: A Research Guide*. Routledge Music Bibliographies Ser. (1st ed., Vol. 56). Routledge Publishing, Inc.

Morgan, A. (2016). Renaissance interval-succession theory: Treatises and analysis. [Doctoral dissertation, McGill University]. Proquest.

Sears, D. R. W., Arzt, A., Frostel, H., Sonnleitner, R., & Widmer, G. (2017, July 18). Modeling harmony with skip-grams. *18th International Society for Music Information Retrieval Conference* (ISMIR), Suzhou, China, 332-338.