The Mozart Expositional Punctuation Corpus: A Dataset of Interthematic Cadences in Mozart's Sonata-Allegro Exposition

OMER RAZ
Tel Aviv University

DROR CHAWIN
Tel Aviv University

URI B. ROM[1]
Tel Aviv University

ABSTRACT: This report documents a dataset consisting of expert annotations (symbolic data) of interthematic (higher-level) cadences in the exposition sections of all of Mozart's instrumental sonata-allegro movements.

INTRODUCTION

MUSICAL punctuation—that is, the segmentation of musical pieces through closural events (typically referred to as "cadences")—is a pivotal phenomenon in all music, and the study of cadences and cadential formulas represents an essential part of the basic training in common-practice tonal music (cf., e.g., Aldwell, Schachter & Cadwallader, 2011; Kostka, Payne & Almén, 2013). Although historical sources discussing punctuation and cadences, such as Koch (1782–1793) and Riepel (1752–1768), were mainly designated for the instruction of beginning composers, they are often called upon in recent analytical literature (e.g., Burstein, 2010; Diergarten & Neuwirth, 2019) to provide historically informed insights into the structure of eighteenth-century music. Cadence identification and analysis in general represent highly topical research objects at the junction between *Formenlehre* (e.g., Rosen, 1988; Caplin, 1998, 2004; Webster, 2001; Hepokoski & Darcy, 2006), contextualized harmonic analysis (e.g., Caplin, 2004; Rohrmeier & Neuwirth, 2015; Martin & Pedneault-Deslauriers, 2015), schema theory (e.g., Temperley, 2004; Gjerdingen, 2007), and computational and empirical musicology (e.g., Ito, 2014; van Kranenburg & Karsdorp, 2014; Duane, 2019). With the expert-annotated dataset introduced here, comprising a sizable core repertoire of the Classical style, and the detailed methodological guidelines accompanying it, we engage in a dialogue with state-of-the-art research on Classical form and the Classical cadence, and contribute to a growing corpus of machine-readable datasets of music-analytical annotations.

METHODS

Dataset

The dataset presented here consists of expert annotations of interthematic cadences (i.e., cadences that punctuate entire themes) in the exposition sections of Mozart's instrumental sonata-allegro movements (with the exception of concerto movements, see Corpus Selection). The cadences are labeled according to the four expositional closure types described by H. C. Koch (see Annotation Standard), and identified by timestamps representing their respective arrival and end positions (whereby a cadence's end is construed to coincide with the next unit's beginning), with additional data on elision and tonal embedding. The dataset is licensed under Creative Commons License (v1.0 Universal) and is hosted with version control in a GitHub repository

(https://github.com/corpusTAU/The-Mozart-Expositional-Punctuation-Corpus) alongside a detailed guidelines report and a user-friendly visualization of the dataset. It is freely available for non-commercial academic, creative, or other uses.

Corpus Selection

Mozart's creative career covers the last third of the eighteenth century and is largely considered to epitomize the late Classical style. His movements in sonata form exemplify a large variety in terms of their scope, genre, and instrumentation, and attest to the stylistic changes that his music underwent from his earliest sonata movements of 1764 through the last year of his life, 1791. The present corpus includes fast movements only, as slow movements in sonata form differ fundamentally in their proportions and internal organization from fast ones (cf. Rosen, 1988). While the sonata-allegro design is most typically associated with opening movements of multi-movement instrumental cycles (such as symphonies and sonatas), this design is also common to many finales (and a few movements located at other positions within the cycle), which are accordingly also included. Concerto movements in sonata form differ crucially in scope and organization from non-concerto movements (cf. Hepokoski & Darcy, 2006), and are accordingly excluded from the corpus (this also refers to concerto-like movements in non-concerto works, such as serenades, etc.). Some hybrid formal designs, such as sonata-rondos (a hybrid between the sonata and the rondo form), marches (a hybrid between the binary and the sonata form), and two sonata-like movements with two-tempo expositions (K. 99/vii and K. 303/i) are excluded as well. We include opera overtures and other sonata-allegro movements from Mozart's stage works, as they share crucial characteristics with Mozart's symphonic sonata-allegro movements (in multi-section overtures, only the sonata-allegro portions are included; in addition, if a given overture exists in a concert-hall version, only the latter version is included). The corpus comprises 298 sonata-allegro movements in a variety of genres (orchestral music, instrumental movements from stage works, ensemble music, chamber music, and music for piano solo and piano four hands).

Expert Annotations

Scores of Mozart's instrumental music were obtained from the authoritative *Neue Mozart Ausgabe* (1955-) in the form of digital scans (publically available online from https://dme.mozarteum.at/). Annotations were manually entered into the score and subsequently collected in the dataset. For a sub-corpus comprising 128 movements (about 43% of the corpus), UR and OR annotated the scores independently of one another; then, UR established the final annotation version presented in the dataset. For the remainder of the corpus, OR entered the annotations, while UR supervised and corrected them where necessary. Further mutual cross-checking was applied to ensure uniformity of the annotation standard across the corpus.

Annotation Standard

Heinrich Christoph Koch, in his extensive composition manual, supplied a particularly comprehensive description of musical punctuation, including detailed observations regarding punctuation in what modern theories refer to as movements in sonata form (1782–1793, vol. 3, pp. 366 ff.). Koch identifies in the first part of such movements—the part termed "exposition" in modern theory—four types of major closural events occurring in a typified order: 1. a (light) authentic cadence in the movement's main key, 2. a half cadence in the main key, 3. a half cadence in the secondary key, and, finally, 4. an authentic cadence in the secondary key. According to Koch, some of these cadences may be omitted or duplicated in a given exposition.

The four closure types described by Koch (or a selection of them) occur in every Classical sonata exposition. Thus, our dataset—comprising over 1,400 Koch-based cadential labels across all of Mozart's instrumental movements of the sonata-allegro type (aside from the omissions referred to above)—yields significant insights into questions of Classical form and possible correlations between the occurrence of cadences and other structural parameters.

Table 1 specifies the four cadential labels used in our corpus alongside basic information regarding the corresponding harmonic degrees (cadential and half-cadential goals) in major- and minor-mode expositions (note that the actual closural sonorities may vary in some unessential details from the ones specified, e.g., due to chord inversions or an added seventh in the half-cadential types). The last column specifies the possibilities of formal embedding of each of the four closure types within the expositional trajectory according to modern theories of sonata form.

Table 1. Tonal areas, arrival harmonies, and formal positions for the four cadence labels.

Label	Cadence attained in key	Degree in major-mode pieces	Degree in minor-mode pieces	Typical formal position(s)
K1	Primary key (I or i)	I	i	End of primary theme (or, alternatively, of a thematic module in the primary-theme zone); end of an incomplete section of the primary theme (in case there is no theme-concluding closure on the principal-key tonic)
K2	Primary key (I or i)	V	V	Launching the transition section or part of it; preceding the secondary theme; alternatively, concluding a sufficiently independent part of the primary theme
K3	Secondary key (V in major; III or, seldom, v in minor)	V:V	III:V (in a single case: v:V)	Heralding the secondary theme (or, alternatively, a transition theme in the secondary key); alternatively, elided or embedded in the transition section (i.e., not realized as a veritable caesura), or preceding a secondary-theme module that occurs at a non-initial position of the secondary-theme zone
K4	Secondary key (V in major; III, or, seldom, v in minor)	V:I	III:I (in a single case: v:i)	End of secondary theme (or of a subordinate thematic module of the secondary-theme zone); end of a closing theme or a codetta; occasionally, concluding the transition section and introducing the secondary theme

Although our annotation standard is based on Koch's system of the four closural types in a sonata exposition, its application to Mozart's music requires crucial adjustments. The labeling of interthematic expositional cadences involves two main levels of decision making: a) identifying valid closural events (authentic cadences and half cadences), b) selecting those cadences that serve as major punctuation events in a given exposition—most typically, such cadences are expected to delineate a thematic module's beginning or end, and are accordingly referred to as interthematic. Detailed criteria for the identification and annotation of interthematic cadences are supplied in the guidelines report published alongside our dataset.

The Mozart Expositional Punctuation Corpus contains the following data for each of the labeled cadences:

GENERAL DATA AT THE WORK LEVEL

The work's identifying number (`<k_id>`), based on the traditional Köchel numbering, its title (`<work>`), full Köchel number (`<k>`), genre (`<genre>`), and year of composition (`<year>`)—the latter field specifies the recorded/conjectured date of completion based on Rom (2011). Compositions in the dataset are ordered by ascending Köchel number; note, however, that this only partly corresponds with the works' chronology.

DATA AT THE MOVEMENT LEVEL

The field `<mov_id>` contains a movement's numeric identifier that is assigned uniquely across the entire corpus. Additional fields specify the movement's ordering number in the work (`<mov>`; in multi-movement works only—otherwise left blank), its main tempo and time signature (`<tempo>`, `<time_sig>`; notably, these parameters remain constant throughout a given exposition section across the entire corpus), and key (`<key>`).

Exposition beginning (`<exp_beg>`): Most commonly, the exposition section begins at m. 1 (or, alternatively, with an upbeat to the first measure); however, slow introductions displace the beginning of the sonata design proper to a later position across the movement.

Exposition length (`<exp_len>`): The length of the exposition section in measures. If there is no repeat sign marking the exposition's end, this value is inferred analytically.

Total cadences (`<tot_cads>`): This field informs about the total number of interthematic cadences annotated in a given exposition.

DATA AT THE INDIVIDUAL CADENCE'S LEVEL

Cadence ordinal number (`<cad_no>`)**:** This field represents a cadence's position within the sequence of annotated interthematic cadences in a given exposition; the last cadence in a given exposition receives the highest value which equals the `<tot_cads>` value for this exposition.

Cadence type (`<cad_type>`): This field informs about the type of a given interthematic closure according to Koch's four types. In our dataset we opt for the labels "K1," "K2," "K3," and "K4," the "K" standing for an abbreviation of "Koch." By using these labels, we propose to emphasize the hierarchical priority of the annotated cadences over merely local cadences that attain the same tonal goals, a distinction that is easily blurred if cadences are designated using their harmonic degrees (cf., for instance, "V of V (HC)" in Burstein (2010), standing for our "K3"). Moreover, our annotation standard makes it possible to use the same labels for movements in major and minor, although the labels "K3" and "K4" denote different harmonic degrees depending on the movement's mode (see Table 1).

Key of subsequent section (`<next_key>`): Whereas a cadence's label implicitly informs about the key in which a specific closure type is attained, the key of the formal unit directly following a given closure is not automatically predictable—this field supplements such missing information. With regard to K1 and K2-type cadences, which are attained in the home key, the subsequent key varies between the principal key (typically following a K1-type cadence) and the secondary key (typical after K2-type cadences). K3 and K4 cadences are—by definition—arrived at and followed by passages in the secondary key (in both major- and minor-mode movements). We leave this field blank with regard to the last annotated closure in an exposition (which is invariably of the K4 type).

Elision with subsequent section (`<elided>`): This Boolean field receives a "true" value under one of the following conditions: a) if the point of cadential arrival coincides with the beginning of the subsequent unit (*direct elision*), or b) in the case of an elided postcadential extensions (*indirect elision*)—see our guidelines report for further details.

Measure and beat of cadential arrival (`<arr_m>`, `<arr_beat>`): This timestamp, consisting of a measure and a beat number, refers to the point of cadential arrival (cf. Caplin, 2004). If a cadence involves melodic suspensions (or a suspended harmony), the point of cadential arrival is considered to be that of the metrically accentuated suspension (rather than that of the resolution). The beat value of cadential arrival (as well as that of the next unit's beginning, see below) starts from 1 (standing for the downbeat), and is represented with respect to the time signature, meaning that in 2/2 time the beat equals a half note, in 2/4, 3/4 and 4/4 time a quarter note, and in 3/8, 6/8 and 12/8 time a dotted quarter note.

Measure and beat of closural event's end (`<end_m>`, `<end_beat>`): We opt for identifying a closural event's end as the beginning timestamp of the next unit. (At an exposition's end, the next event is either the development section's beginning, or a retransition passage that nominally still belongs to the exposition, but is located past the final closure in the secondary key.) With reference to Caplin (2004), we term the space between the cadential arrival and the next unit's beginning *postcadential space* (in the case of direct elision, this space has a size of zero). We define postcadential space as consisting of the following optional constituents (see also Figure G1 in our guidelines report): (1) a *postcadential extension* consisting of the closural degree itself (unless directly elided, in which case there is no postcadential extension), and

any subsequent progressions that reiterate or reestablish it; (2) an optional *caesura*, most typically a simultaneous rest in all parts; (3) an optional *postcadential link* connecting to the subsequent new beginning (we construe as links any athematic connecting passages, as well as upbeat figures shorter than one measure that thematically belong to the following unit). In the case of *direct* elision, the timestamp of a closural event's end is identical with that of cadential arrival, and in the case of *indirect* elision, it is identical with that of the postcadential extension's end (see our guidelines report for details). If the next beginning is construed to start off-beat, the beat value of the closure's end will be indicated by a fraction expressing the new beginning's position relative to the beat length (a ⅓ is rounded down to .33).

> **Closural event's length** (`<cad_len>`): This field corresponds to the length of the postcadential space which is delineated, on the one side, by the cadential arrival, and, on the other, by the closural event's end (see discussion above and guidelines report). Values are given in quarter notes, meaning that, for instance, in 6/8 time fractions may occur also where both the arrival and the end timestamps fall on the beat. Although this value can be inferred from the arrival and end timestamps described above (assuming knowledge of the beat's length), we nevertheless include it to facilitate calculations. In case of direct elision at the point of cadential arrival, this field has a zero value. Otherwise, we construe the postcadential space to be of regular size if it spans one measure in common time (4/4), *alla breve* (2/2), 3/4, 6/8 and 12/8 time. In 2/4 and 3/8 time, we allow for two measures. We consider any size above these values to represent an *augmented postcadential space*.

> Figure 1, taken from the first movement of Mozart's Sonata for Two Pianos K. 448, provides a demonstration of Mozart's complex use of expositional cadences. The passage in question encompasses four closural events of the four types K1–4. The K1-type cadence in m. 17 concludes an extensive presentation of the movement's compound-sentential primary theme. It is directly elided with the subsequent unit which displays typical transition characteristics (quick rhythmic values, sequential harmonic blocks). The half cadence in m. 24 strikes the listener as an interthematic closure potentially preparing the onset of the secondary theme; however, this theme-introducing potential remains unrealized, as the closure is elided with a one-measure passage which modulates from D major to the secondary key, A major, using the latter key's raised fourth scale degree (D♯) to introduce yet another half cadence in m. 25, this time of the K3 type. This latter cadence receives an extensive postcadential extension, which re-arrives at the closural degree at m. 30 (again, employing the secondary key's raised fourth scale degree, D♯, in m. 29); however, at this point the secondary-key dominant is abandoned, and the music heads towards the secondary-key tonic A major, with cadential arrival at m. 33. According to our analysis, the passage following the K3-type cadential arrival at m. 25 consists of a postcadential extension stretching through the downbeat of m. 30, at which point it is elided with the beginning of a further short transitory passage encompassing mm. 30–33 (while this latter passage resembles a postcadential link, we do not construe is as such, as it isn't merged with the following unit, but rather closed by a clear punctuation at m. 33). Hence, the K3-type cadential event beginning at m. 25 is considered as ending on the downbeat of m. 30, with a total length of five measures in common time (20 quarter notes)—a constellation indicative of an *augmented postcadential space*. Following the two elided half cadences at mm. 24 and 25–30—both of which may be construed as declined medial caesuras (cf. Hepokoski & Darcy, 2006, pp. 45 ff.)—the last bit of this complex transition (mm. 30–33) finally attains a veritable theme-introducing caesura of the K4 type announcing the onset of the movement's secondary theme at mm. 34 ff.

SUMMARY STATISTICS

The Mozart Expositional Punctuation Corpus contains expert annotations of higher-level (interthematic) cadences across the exposition sections of all instrumental sonata-allegro movements by Mozart (with certain exceptions as specified above). The analyzed repertoire includes 298 sonata expositions in a variety of genres, with 277 movements (about 93%) in the major and 21 movements (about 7%) in the minor mode. The annotation corpus comprises 1,450 annotations using a system of four cadence-type labels ("K1," "K2," "K3," and "K4") complemented by additional data as described above. In total, the exposition sections analyzed amount to 19,173 measures of music. Expositions' lengths average at 64.3 measures (disregarding differences of time signature among expositions) with median at 56.5 measures. Individual exposition lengths range from a minimum of 9 to a maximum of 212 measures (the latter value refers to a movement in 2/4 time), with a total of two to ten interthematic cadences per exposition, yielding an average of 4.9 interthematic cadences per sonata exposition. (The maximal number of cadences, 10, occurs in a single movement: the opening movement of the *Jupiter* Symphony.)

Figure 1. W. A. Mozart, Sonata for Two Pianos in D major, K. 448/i, mm. 16–33 with closural labels.

As indicated in Table 2, some of the expositions in the corpus contain multiple interthematic cadences of the same type (either in direct succession or in irregular order), with up to six K4 labels in a single exposition.

Table 2. Summary statistics of some major properties of the corpus.

	K1	*K2*	*K3*	*K4*	*All types*
Total number of label occurrences in corpus (298 movements)	300 (20.7%)	240 (16.6%)	207 (14.3%)	703 (48.5%)	1450 (100%)
Minimal number of occurrences in an exposition	0	0	0	1	
Maximal number of occurrences in an exposition	2	3	2	6	
Maximal number of consecutive occurrences in an exposition	2	3	2	6	
Total number of occurrences in 277 major-mode movements	280	227	190	654	1351 (93.2%)
-following section begins in the tonic major (I)	270	17	0	0	287
-following section begins in the dominant key (V)	10	210	190	377	787
Major-mode movements with no instances of closure type	8	57	97	0	
Total number of occurrences in 21 minor-mode movements	20	13	17	49	99 (6.8%)
-following section begins in the tonic minor (i)	18	5	0	0	23
-following section begins in the relative major key (III)	2	8	16	28	54
-following section begins in the dominant minor (v)	0	0	1	0	1
Minor-mode movements with no instances of closure type	3	8	7	0	
Closure elided with next unit	143	61	20	273	497 (34.3%)
Instances with augmented postcadential space	57	72	112	148	389 (26.8%)
Average size of postcadential space (in quarter notes)	4.9	6.6	12.4	4.0	5.8

The table shows the distribution of labels across the corpus, with K4-type cadences representing about 48.5% of all labels. K4 is also the only mandatory closure type, occurring in all expositions. Whereas only a few expositions lack a K1-type closure, the half-cadential types K2 and K3 are quite often missing. (Additional analysis shows that only seven expositions lack both K2 and K3; on the other hand, a co-occurrence of both labels is quite common, and is found in 136 expositions.) Also shown in the table is the labels' distribution across major- and minor-mode expositions, further differentiated by the subsequent key. Over a third of all cadences (and almost 48% of all K1-type cadences) are elided with the subsequent unit's beginning. Almost 27% of all cadential arrivals (and over a half of all K3-type cadences) are followed by an augmented postcadential space (see definition above). Calculated across the entire corpus, the average size of the postcadential space is nearly 6 quarter notes, with K3-type cadences displaying the highest average value (above 12 quarter notes), and peaking at 17 measures in 4/4 time (in the first movement of the String Quintet in C major, K. 515, mm. 69–85).

As shown in Table 3 and Figure 2, the movement count peaks at 5 interthematic cadences per exposition, with 98 expositions (almost a third of the corpus). Figure 3 shows that, while there is a great variability in exposition lengths across the corpus, the number of interthematic cadences per exposition correlates positively with expositions' lengths, as expected.

Table 3. Movement counts and average/median exposition lengths (in measures) by total number of interthematic cadences per exposition.

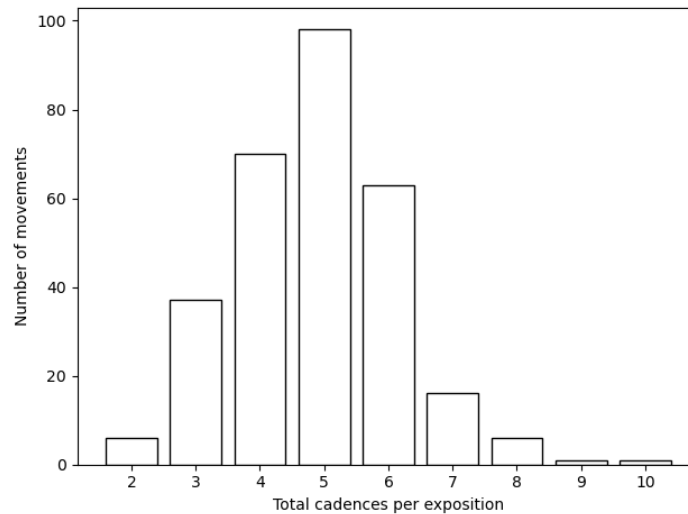|	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Count of movements	6	37	70	98	63	16	6	1	1
Average exposition length	31.8	42.5	50.3	66.4	82.4	79.0	115.5	106.0	120.0
Median exposition length	33.5	42	44.5	58	78	78	101.5	106	120

Figure 2. Movement count as a corollary of the total number of interthematic cadences per exposition
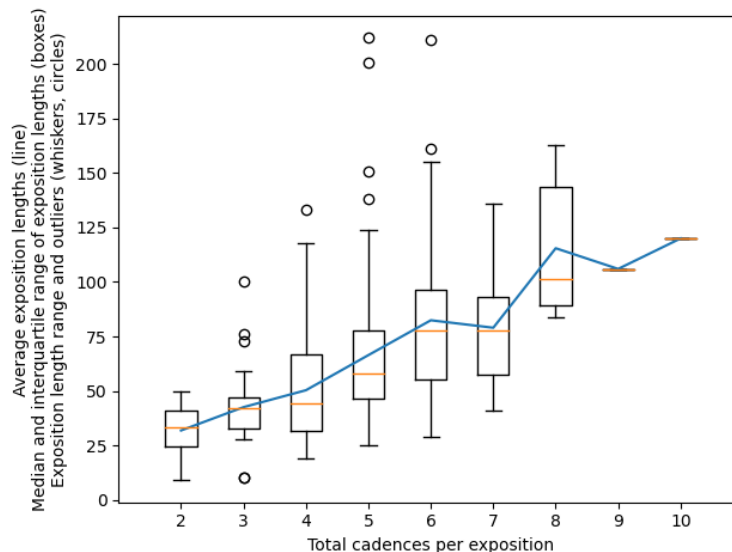
Figure 3. Average exposition length as a corollary of the total number of interthematic cadences per exposition (line), median and interquartile range of exposition lengths (boxes), range of exposition lengths and outliers (whiskers and circles).

DATA USAGE AND APPLICATIONS

Cadence detection is a highly topical research object of quantitative and computational musicology, and represents a crucial part of a broader effort to model musical form using data-driven methods (see, e.g., Allegraud et al., 2019; Ito, 2014; Sears, Pearce, Caplin & McAdams, 2018). Whereas musical corpus studies may now draw on an ever-growing body of datasets of harmonic analyses, explicit cadential labels are, by and large, rare. (Consider, for instance, the case of the recently published Annotated Beethoven Corpus (Neuwirth, Harasim, Moss & Rohrmeier, 2018), whereby closural events can be indirectly inferred from end-of-phrase markers.) Expert annotations, being an essentially subjective matter (cf. Koops et al., 2019), call for comparing analyses from various sources across the music-analytical community. We expect that by juxtaposing our cadential labels as ground-truth data with other annotation corpora as well as symbolic and audio representations of the works analyzed, our dataset will contribute toward an ever deeper and more comprehensive understanding of the complex interrelations among harmonic, contrapuntal and formal procedures in the music of the late eighteenth century. The carefully developed and comprehensively documented methodological guidelines accompanying our dataset publication engage in a dialogue with state-of-the-art analytical literature on Classical form and the Classical cadence: through this discussion, we seek to flesh out the theoretical perspectives of corpus analysis for a possible rethinking of some of the current analytical views on the Classical sonata form and cadence distribution across Classical sonata expositions.

AUTHOR CONTRIBUTIONS

OR and UR annotated the corpus and were involved in equal parts in the development of the guidelines and the annotation standard. DC programmed the project database and supplied the summary statistics. UR conceived the project and wrote the article in collaboration with OR and DC. This research was supported by The Israel Science Foundation Grant No. 480/17 to UR.

ACKNOWLEDGMENTS

NOTES

[1] Correspondence can be addressed to: Uri Rom, PhD, The Buchmann-Mehta School of Music, Tel Aviv University, P.O.B. 39040, Tel Aviv 6997801, Israel, urom@tauex.tau.ac.il.

REFERENCES

Aldwell, E., Schachter, C., & Cadwallader, A. (2011). *Harmony and Voice Leading* (4th ed.). Beverly, MA: Wadsworth Publishing.

Allegraud, P., Bigo, L., Feisthauer, L., Giraud, M., Groult, R., Leguy, E., & Levé, F. (2019). Learning Sonata Form Structure on Mozart's String Quartets. *Transactions of the International Society for Music Information Retrieval*, *2*(1), 82–96. https://doi.org/10.5334/tismir.27

Burstein, P. (2010). Mid-Section Cadences in Haydn's Sonata-Form Movements. *Studia Musicologica*, *51* (1/2), 91–107. https://doi.org/10.1556/smus.51.2010.1-2.7

Caplin, W. E. (1998). *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. New York, NY: Oxford University Press.

Caplin, W. E. (2004). The Classical Cadence: Conceptions and Misconceptions. *Journal of the American Musicological Society*, *57*(1), 51–117. https://doi.org/10.1525/jams.2004.57.1.51

Diergarten, F., & Neuwirth, M. (2019). *Formenlehre: Ein Lese- und Arbeitsbuch zur Instrumentalmusik des 18. und 19. Jahrhunderts*, Laaber, Germany: Laaber-Verlag.

Duane, B. (2019). Melodic Patterns and Tonal Cadences: Bayesian Learning of Cadential Categories from Contrapuntal Information. *Journal of New Music Research*, *48*(3), 197–216. https://doi.org/10.1080/09298215.2019.1607396

Gjerdingen, R. O. (2007). *Music in the Galant Style*. New York, NY: Oxford University Press.

Hepokoski, J., & Darcy, W. (2006). *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth-Century Sonata*. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195146400.001.0001

Ito, J. P. (2014). Koch's Metrical Theory and Mozart's Music: A Corpus Study. *Music Perception: An Interdisciplinary Journal*, *31* (3), 205–222. https://doi.org/10.1525/mp.2014.31.3.205

Koch, H. C. (1782–1793). *Versuch einer Anleitung zur Composition*, 3 vols., Leipzig: Bohme.

Koops, H. V., de Haas, W. B., Burgoyne, J. A., Bransen, J., Kent-Muller, A., & Volk, A. (2019). Annotator Subjectivity in Harmony Annotations of Popular Music. *Journal of New Music Research*, *48*(3), 232–252. https://doi.org/10.1080/09298215.2019.1613436

Kostka, S., Payne, D., & Almén, B. (2013). *Tonal Harmony, with an Introduction to Twentieth-Century Music* (7th ed.). New York, NY: McGraw-Hill.

Martin, N. J., & Pedneault-Deslauriers, J. (2015). The Mozartean Half Cadence. In M. Neuwirth & P. Bergé (Eds.), *What Is a Cadence? Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire* (pp. 185–213). Leuven, Belgium: Leuven University Press. https://doi.org/10.2307/j.ctt14jxt45.9

Mozart, W. A. (1955-). *Neue Mozart Ausgabe. Neue Ausgabe sämtlicher Werke*. Kassel, Germany: Bärenreiter. Retrieved from International Foundation Mozarteum, https://dme.mozarteum.at/, Salzburg, Austria.

Neuwirth M., & Bergé, P. (Eds.). (2015). *What Is a Cadence? Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire*. Leuven, Belgium: Leuven University Press. https://doi.org/10.11116/9789461661739

Neuwirth, M., Harasim, D., Moss, F. C., & Rohrmeier, M. (2018). The Annotated Beethoven Corpus (ABC): A Dataset of Harmonic Analyses of All Beethoven String Quartets. *Frontiers in Digital Humanities*, *5*. https://doi.org/10.3389/fdigh.2018.00016

Riepel, J. (1752–1768), *Anfangsgründe zur musikalischen Setzkunst* (Vols. 1–5). Augsburg, Germany: J.J. Lotter.

Rohrmeier, M., & Neuwirth, M. (2015). Towards a Syntax of the Classical Cadence. In M. Neuwirth & P. Bergé (Eds.), *What Is a Cadence? Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire* (pp. 285–336). Leuven, Belgium: Leuven University Press. https://doi.org/10.2307/j.ctt14jxt45

Rom, U. (2011). *Tonartbezogenes Denken in Mozarts Werken unter besonderer Berücksichtigung des Instrumentalwerks*, Ph.D. Dissertation. Retrieved October 1, 2020 from: http://opus4.kobv.de/opus4-tuberlin/frontdoor/index/index/docId/2988.

Rosen, C. (1988). *Sonata Forms*. New York, NY: W.W. Norton.

Schoenberg, A. (1967). *Fundamentals of Musical Composition* (G. Strang and L. Stein, Eds.). London, UK-Boston, MA: Faber & Faber.

Sears, D. R., Pearce, M. T., Caplin, W. E., & McAdams, S. (2018). Simulating Melodic and Harmonic Expectations for Tonal Cadences Using Probabilistic Models. *Journal of New Music Research*, *47*(1), 29–52. https://doi.org/10.1080/09298215.2017.1367010

Temperley, D. (2004). *The Cognition of Basic Musical Structures*. Cambridge, MA: The MIT Press.

van Kranenburg, P., & Karsdorp, F. (2014). Cadence Detection in Western Traditional Stanzaic Songs using Melodic and Textual Features. In H.-M. Wang, Y.-H. Yang, & J.H. Lee (Eds). *Proceedings of the 15th International Society for Music Information Retrieval Conference* (pp. 391–396). Taipei, Taiwan: ISMIR. https://doi.org/10.5281/zenodo.1416172

Webster, J. (2001). Sonata Form. *Grove Music Online*. https://doi.org/10.1093/gmo/9781561592630.article.26197

ERRATUM

3/14/2022: Corrected GitHub link to https://github.com/corpusTAU/The-Mozart-Expositional-Punctuation-Corpus