Commentary on "The Mozart Expositional Punctuation Corpus: A Dataset of Interthematic Cadences in Mozart's Sonata-Allegro Expositions"

BEN DUANE[1]
Department of Music, Washington University

THE Mozart Expositional Punctuation Corpus is a welcome contribution not only because it is large and soundly analyzed but also because these analyses draw on a theoretical framework not especially in vogue at this moment. While much recent work on sonata-form cadences, including my own, draws mainly on modern theories such as Caplin's (1998) and Hepokoski and Darcy's (2006), Raz et al. (2021) reach back to the eighteenth-century writings of H. C. Koch, providing the sort of differing perspective that is essential to corpus-based research. An inescapable challenge of corpus methods is disentangling the results of studies from the music theories that informed those studies. On the one hand, it is all but impossible to conduct a meaningful corpus analysis without espousing at least a minimal theory of musical structure. On the other hand, the particulars of a study's guiding theory can influence its results and building the same study around a different theory could alter its conclusions. The only solution to this methodological problem is for corpus-based music theory to accumulate datasets that stem from a wide variety of theoretical orientations. Empirical conclusions supported by most or all of this data would earn the mark of converging evidence. Those supported by some datasets but not others would be more suspect. Herein lies a major benefit of this corpus.

Indeed, as I began working with the data, I quickly realized it could throw light on an issue addressed by sonata theories of many stripes: the deployment and timing of authentic cadences toward the ends of expositions. Mozart's expositions frequently end with a series of authentic cadences of the type shown in Figure 1. Such multi-cadence successions have prompted much debate about which cadence marks the end of the second theme and the beginning of closing space. Many theorists argue that the first PAC followed by new material—what Hepokoski and Darcy (2006) call the point of *essential expositional closure*—forms this boundary. Caplin (1998) and others, however, identify the second theme's ending at the last PAC preceded by material that passes for thematic rather than purely closural. (For a fuller account of this debate, see Hepokoski & Darcy, 2006, pp. 120–124).) With regard to Figure 1, this disagreement amounts to asking whether the second theme ends at mm. 39 or 59.

Yet such discussions can belie a larger point about closure in sonata expositions. In expositions ending with many PACs, these cadences effect a sense of closure together that none does individually. Mozart's exposition would not have sounded complete if only the cadence in mm. 39 or 59 were included. Only with the series of cadences including mm. 39 and 59 (and extending through m. 63) does Mozart create an ending equal to the proportions of the exposition. Closure, in this exposition and many others, is a process, not an event.

A common feature of such exposition-ending sequences of PACs is the arrival of cadences at ever-shorter time intervals. In Figure 1, for instance, the first and second PACs are separated by eight bars while subsequent PACs are divided by four or five bars. This acceleration manifests itself not only in cadential arrivals but also in the beginnings of cadential progressions, which, as Figure 1 illustrates, occur sooner and sooner after each prior cadence. The same, incidentally, is true of the exposition in the first movement of Mozart's sonata for two pianos, K. 448, which the authors use as an example of other cadence-analysis issues. In this exposition, the abandonment of the initial second-theme material is followed by five PACs whose arrivals are spaced by 15, 9, 5, and then 1 bars (with the cadential progressions starting after 13, 5, 2, and 1 bars). These analyses, I suspect, would surprise few who are acquainted with Mozart's sonata style. Many more of his first-movement expositions end with this sort of accelerating series of authentic cadences.

Fig 1. Mozart, Piano Sonata in B, K. 333, Mvt. I, mm. 36-63. Annotations show all perfect authentic cadences following the subordinate theme's onset, the progressions that effect these cadences, the four Kochian K4 cadences identified by Raz et al. (2021), the point of *essential expositional closure* that Hepokoski & Darcy (2006) would likely identify, and the ending of the second theme as Caplin (1998) would likely analyze it.

Here is a hypothesis which Raz et al.'s corpus can help evaluate. While not all cadences of the type marked in Figure 1 qualify as one of the four Kochian types identified in the corpus, the K4 type includes many of the PACs that follow the onset of the subordinate theme. These K4 cadences, moreover, also exhibit accelerated, ever-closer placement in the first movements of both K. 333 (mm. 30, 50, 59, 63) and K. 448 (mm. 33, 49, 73, 79). To determine whether this tendency is present across the corpus, I selected the subset of movements that include two or more K4 cadences. For each of these movements, I then computed the correlation between the ordinal number of the K4 cadences and the number of bars that follow the cadences

before the next K4 cadence (or, in the case of each exposition's final K4 cadence, the development's first bar). Under this analysis, negative correlations indicate decreasing time intervals between subsequent cadences, and such negative correlations are obtained in 228 of the 232 movements in the subset ($p < .001$ by a sign test).

In other words, the Mozart Expositional Punctuation Corpus provides evidence supporting the notion that accelerating, post-subordinate-theme cadences are a hallmark of Mozart's sonata expositions. Granted, this is hardly definitive evidence, shaped as it is by a theory of cadences that privileges some cadences over others. But if combined with other results derived from corpora analyzed using other theories, it might contribute to the sort of converging evidence described above. And it might help form the basis of a more detailed empirical analysis of Mozart's use of cadences at the ends of expositions—one that, for example, assesses time intervals not just between one cadence and the next but between one cadence and the next cadential progression.

Which brings me to the first of two small criticisms of this very useful dataset. The details included about both the individual cadences and the movements containing them leave some to be desired. The starting measure of the progression leading to each Kochian cadence would have been useful to include. So would more specifics about each cadence's harmonic content and rhythm, textural and dynamic profile, use of trills and other embellishments, and whether the cadence immediately repeats an earlier one (either completed or evaded). In many ways, however, this is less a comment on the authors' decisions than on a lack of consensus among researchers about what information is worth including in public corpora. This issue warrants ongoing discussion.

My second criticism is that the dataset's formatting does not always make for easy analysis, especially across a variety of software. In particular, the coding of data in many different formats—text strings, Booleans, Roman numerals, integers, integers pairs separated by slashes—complicates the importation and analysis of the data in command-line environments like *R* and *Matlab*, since text expressions such as "K1" and "4/4" must be parsed and translated into numerical values. The same is true for *SPSS* and other spreadsheet-like statistics software, which often recognize only numerical data. Formatting the data with integers whenever possible—replacing *K1* and *K2* with *1* and *2*, *TRUE* and *FALSE* with 1 and 0, *I* and *V* with 1 and 5, and *3/4* with *3* in one column and *4* in another—would have made the data easier to use in a variety of settings.

Yet these are both small shortcomings, vastly outweighed by the value offered by the authors' generous contribution. After a few hours with the data, I uncovered the interesting trends reported above. I suspect that future research will show there is much more to find where these results came from.

ACKNOWLEDGMENTS

NOTES

[1] Correspondence can be addressed to: Ben Duane, PhD, Department of Music, Washington University, CB 1032, One Brookings Drive, St. Louis, bduane@wustl.edu.

REFERENCES

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.

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

Raz, O., Chawin, D., & Rom, U. B. (2021). The Mozart expositional punctuation corpus: A dataset of interthematic cadences in Mozart's sonata-allegro exposition. *Empirical Musicology Review 16*(1), 134–144. https://doi.org/10.18061/emr.v16i1.7648