

The Medial Caesura and Its Role in the Eighteenth-Century Sonata Exposition

James Hepokoski and Warren Darcy

I. STRUCTURAL PUNCTUATION AND SONATA THEORY

As Heinrich Christoph Koch emphasized in the Second Part of his *Versuch einer Anleitung zur Composition* (1787), varying degrees of rhetorical articulation—especially hierarchically ordered cadences, pauses, and breaks—are central to the mid- and late-eighteenth-century sense of form. He referred to this as the principle of “melodic punctuation” (*die melodische Interpunction*) and noted that it was operative on both local and larger structural levels. This foundational idea was introduced in Section 3, “On the Nature of Melodic Parts” (*Von der Beschaffenheit der melodischen Theile*):

By means of these more or less noticeable resting points of the spirit [*Ruhepunkte des Geistes*], the products of [the] fine arts can be broken up into smaller and larger sections . . . Just as in speech, the melody of a composition can be broken up into periods by means of analogous resting points, and these, again, into single phrases [*Sätze*] and melodic parts [*Theile*].¹

We thank William Rothstein, Patrick McCreless, and David Grayson for their helpful comments on an earlier draft of this essay.

¹Heinrich Christoph Koch, *Versuch einer Anleitung zur Composition*, Part 2 (Leipzig, 1787; facsimile ed., Hildesheim: Georg Olms, 1969), 342–43. The translation above is adapted from that of Nancy Kovaleff Baker in Koch, *Introductory Essay on Composition: The Mechanical Rules of Melody, Sections 3 and 4* (New Haven: Yale University Press, 1983), 1. For “*melodische*

As readers of Koch soon learn, he intended his term “melody” to be extended to encompass even an entire movement: broader structural principles are smaller ones writ large. “Melodic punctuation” grows into what we might call “structural punctuation.”

It is our contention that an analysis of major punctuation-breaks (structural caesuras) leads one into the heart of a productive, defensible sonata-form theory. Such a claim is not new. Musical scholars, particularly those informed by the work of Leonard Ratner and others, have been familiar with the idea for decades: those scholars differ only in the degree to which they emphasize this feature in constructing an analytical model. Karol Berger has recently made several strong statements on its behalf in discussions of Koch and the primacy of “punctuation form” in the late-eighteenth-century style; Charles Rosen had acknowledged a similar point about “breaks in texture” and cadence-placement in *Sonata Forms*.²

Interpunction,” see Koch, *Versuch*, 345 (Section 78), and Koch, *Introductory Essay*, 2.

²Karol Berger, “Toward a History of Hearing: The Classic Concerto, A Sample Case,” *Convention in Eighteenth- and Nineteenth-Century Music: Essays in Honor of Leonard G. Ratner*, ed. Wye J. Allanbrook, Janet M. Levy, and William Mahrt (Stuyvesant, N. Y.: Pendragon, 1992), 405–29; “The First-Movement Punctuation Form in Mozart’s Piano Concertos,” in *Mozart’s Piano Concertos: Text, Context, Interpretation*, ed. Neal Zaslaw (Ann Arbor:

Nor has current music theory been silent on this issue. In a recent inquiry into late-eighteenth-century structures (including sonata form) from a Schenkerian perspective, William Rothstein began with fundamental principles:

For Koch, form is [closely] identified with cadence structure: according to his definition, only a section of music ending with a perfect authentic cadence may be termed a “period.”

As Ratner rightly emphasizes, the 18th-century view of sonata form stresses cadences—the ends of sections—rather than thematic statements, which generally sound like beginnings (partly because they tend to occur just after important cadences) . . . This emphasis on cadential goals matches perfectly Schenker’s conception of sonata form.³

So much might seem self-evident, common-sense itself. Nevertheless, once one decides to develop these observations into an expanded sonata theory, complications, individual realizations, variants, and exceptions multiply rapidly. The present authors have been developing such a theory, in large part generated inductively from the analysis of hundreds of sonatas, symphonies, overtures, quartets, and other chamber music from the late eighteenth- and early nineteenth-century repertory. The works examined include not only those of Haydn, Mozart, and Beethoven, but also those of other composers preceding and surrounding them, including Sammartini, Stamitz, Cannabich, J. C. Bach, C. P. E. Bach, Dittersdorf, Boccherini, Clementi, Dussek, Cherubini, and others.

Our goal has been to seek a deepened understanding of the sonata as a historical genre. More specifically, we have been trying to construct a more adequate description of the event-zones (primary-theme zone, transition, medial caesura,

secondary-theme zone, closing zone, and so on) within normative sonata construction in different times and in different places. Our most fundamental conclusion is this: *each event-zone within the sonata-genre is describable as a family of hierarchically ordered standard options available to the composer, analogous to a menu of formatting options within a computer program.* Moreover, at any point within sonata construction, a composer may choose either to realize a standard option in a more or less straightforward manner (thus reaffirming the norm) or to treat that option more flamboyantly—perhaps even submitting it to a *deformation* (stretching it to or even beyond its limits) or overriding the norm altogether for a particular expressive effect. It is precisely in such a personalized treatment of otherwise conventional features that the distinctive style and brilliance of a Haydn, a Mozart, or a Beethoven lies.

In our view, moment-to-moment compositional choices may be profitably understood as elements of an ongoing dialogue with reasonably ascertainable, flexible generic norms. Trying to reconstruct this tacit dialogue can reawaken the expressive power of a piece in ways that we believe to be remarkable.⁴ Our aim has been twofold: first, to (re-)generate those norms—again, inductively (under the conviction that the most valuable treatises on late-eighteenth-century “sonata form” were written by the great masters, not by the early theorists); second, to configure the norms into an ordered description of standard practices, deformations, and overrides that we call “Sonata Theory” (with capital S and

⁴This is a different proposition from that provided by the usual “conformational” theories of the sonata, which have typically erred through overstatement and overdefinition. An enlightening discussion of the historically competing scholarly views of the sonata (“generative” versus “conformational” views, “harmonic” or “linear-contrapuntal” versus “melodic” views, and so on), along with a sensible conclusion for current practice, appears in Mark Evan Bonds, *Wordless Rhetoric: Musical Form and the Metaphor of the Oration* (Cambridge: Harvard University Press, 1991), 1–52.

University of Michigan, 1996), 239–59. Charles Rosen, *Sonata Forms*, rev. ed. (New York: W. W. Norton, 1988), 99.

³William Rothstein, *Phrase Rhythm in Tonal Music* (New York: Schirmer, 1989), 111–13.

T). Sonata Theory seeks to be compatible with current music-theoretical (for example, Schenkerian) and established musicological approaches to the sonata. We present our conclusions thus far as cautious propositions, not as unalterable claims. We invite further refinement and/or correction of these conclusions.

Genre theory is a difficult matter, and we do not wish to minimize its complexities. We realize the gravity of the questions that any theory of genre involves: the problem of defining the concept itself; vexing issues in the ontology of genres, including production and reception genres; the simultaneous existence of multiple, individualized conceptions of the genre (was Haydn's understanding of "sonata form," say, in 1785, the same as Mozart's?); the fluid nature of genres, which change and transform throughout historical time; the thorny matter of the horizon of expectations; and so on. We intend to address these conceptual matters in a separate essay. For now, it has seemed best to provide examples of the practical results of our work and to begin with an examination of one of the linchpins of Sonata Theory: the issue of structural punctuation.

2. THE TWO EXPOSITION TYPES

Within eighteenth-century sonata form, differing treatment of cadences and caesuras in individual compositions produced strikingly varied rhetorical shapes. In turn, these multiple shapes may be grouped into families, each sharing certain crucial features. The differing rhetorical patterns produced within the sonata exposition, for example, may be regarded as falling into two broad families, or *exposition types*. We refer to them as the *two-part exposition* and the *continuous exposition*.

The first type, the *two-part exposition*, was by far the most normative model of the late eighteenth century. As elabo-

rated in this essay, the two-part exposition is characterized by a strong mid-expositional punctuation break, the *medial caesura*—most often articulating a half cadence—followed (almost invariably) by a rhetorical drop to *piano* marking the onset of a gentle, usually contrasting *secondary-theme zone* in the second key-area. This exposition type is preponderant in Mozart and early Beethoven, and it became not only the most significant format for nineteenth-century expositions (sometimes in transformed or deformational variants) but also the model passed on by textbooks—to the exclusion of the continuous-exposition alternative. This was the exposition type alluded to by Riepel in 1755 and Vogler in 1778; described by Koch in 1793 (in which the first theme can be “ein erweiterter, oder mit mehr melodischen Theilen verbundener, und etwas rauschender Satz” [“a somewhat rushing passage, expanded or connected with more melodic parts”]) and the second, in the dominant, is usually “ein cantabler Satz” [“a *cantabile* passage”]), by Galeazzi in 1796 (in which the second theme is the “passo caratteristico” [“characteristic passage”]), and by Kollmann in 1799 (in which the second theme is the “second subsection,” also identified as “a first sort of elaboration,” occupying the latter half of the first “section” of a typical “long movement”). It was also the exposition type outlined in Reicha's famous 1824–26 diagram of “la grande coupe binaire” (“the grand binary design,” whose “première partie, ou exposition des idées” included a “première idée mère” [literally, “first mother-idea”], a “pont” [“bridge”], and a “seconde idée mère dans la nouvelle tonique” [“second mother-idea in the new key”]), and it was most notably elaborated after Reicha by Birnbach in his 1827 outline of what he called the “Hauptform eines grössen Tonstückes” (the “principal form of a large composition,” including an initial “Thema” followed by “der zweite Gedanke” [“second thought”] or “das zweite Thema”—the first use of the specific term “second theme”), Gathy in 1835 (“Hauptgedanke” and “Nebengedanke”), Czerny in c. 1837–49

(“principal subject” and “middle subject”), Marx in 1837–47 (“Hauptsatz” and “Seitensatz”), and others. In addition, in some of his sketches from the years around 1800 (the op. 18 quartets, the violin sonatas from op. 30, the piano sonatas from op. 31), Beethoven seems occasionally to have referred to this “second theme” with the abbreviation, “m.g.” As William Drabkin has speculated, this (foreshadowing the terminology of Czerny?) may have been an abbreviation for “Mittel-Gedanke” (“middle thought”).⁵

The second exposition type, the *continuous exposition*, is encountered frequently in works of the second third of the eighteenth century and in several of the works of Haydn, who

⁵Early descriptions of the sonata are documented in many sources: e.g., Fred Ritzel, *Die Entwicklung der 'Sonatenform' im musiktheoretischen Schrifttum des 18. and 19. Jahrhunderts* (Wiesbaden: Breitkopf & Härtel, 1968); Birgitte Moyer, “Concepts of Musical Form in the Nineteenth Century with Special Reference to A. B. Marx and Sonata Form” (Ph.D. diss., Stanford, 1969); William S. Newman, *The Sonata in the Classic Era: The Second Volume of A History of the Sonata Idea*, 2nd ed. (New York: W. W. Norton, 1972), 19–42; Leonard G. Ratner, *Classic Music: Expression, Form, and Style* (New York: Schirmer, 1980), 217–47; Ian Bent, *Analysis* (New York: W. W. Norton, 1987), 12–32. The above refer to virtually all of these figures. More specifically: for Koch, see note 1 above (in this case the quotations are taken from Part 3 of the *Versuch*, 1793, Sections 141 and 147, 364, 385; cf. the slightly differing translation—using “phrase” for “Satz” where we prefer “passage”—in Baker, *Introductory Essay*, 221, 230). Bathia Churgin discusses Galeazzi in “Francesco Galeazzi’s Description (1796) of Sonata Form,” *Journal of the American Musicological Society* 21 (1968): 181–99. For Kollmann, see *An Essay on Practical Musical Composition* (facsimile ed., New York: Da Capo, 1973), 5–6, and Ratner, *Classic Music*, 219. For Reicha, see Bent, *Analysis*, 18–20, and especially Peter A. Hoyt, “The Concept of *développement* in the Early Nineteenth Century,” in *Music Theory in the Age of Romanticism*, ed. Ian Bent (Cambridge: Cambridge University Press, 1996), 141–62. For Birnbach and Gathy, see Moyer, “Concepts of Musical Form,” 56–57, and Bent, *Analysis*, 25. For Czerny and Marx, see Moyer, “Concepts of Musical Form,” 65 and 69–125. For Beethoven’s “m.g.,” see William Drabkin, “Beethoven’s Understanding of ‘Sonata Form’: The Evidence of the Sketchbooks,” in *Beethoven’s Compositional Process*, ed. William Kinderman (Lincoln: University of Nebraska Press, 1991), 14–19.

employed it throughout his career. It has been only cursorily treated, however, in the scholarly literature.⁶ The foremost characteristic of the continuous exposition is the absence of the defining feature of a two-part exposition: a medial caesura that makes possible a subsequent, conceptually separable secondary-theme zone.

Although an adequate discussion of continuous expositions must await a separate study, we should mention that several subtypes may be discerned, of which two are particularly noteworthy. In the first, more common subtype, a brief primary-theme zone or initial thematic module precedes a relentlessly ongoing, broadly elaborated *Fortspinnung*—an “expansion section” (*Entwicklungspartie*), as it has been called—that occupies the bulk of the exposition and concomitantly avoids (or evades) structural caesuras (including decisive perfect authentic cadences) at or around the mid-

⁶Apparently the first scholar to identify this “unusual” form was Jens Peter Larsen, who in 1963 called it a “three-part division of the exposition” (*Dreiteilung . . . der Exposition*) in “Sonata Form Problems,” *Handel, Haydn, and the Viennese Classical Style*, trans. Ulrich Krämer (Ann Arbor: UMI, 1988), 269–79; orig. publ. as “Sonatenform-Probleme,” in *Festschrift Friedrich Blume zum 70 Geburtstag*, ed. Anna Amalie Abert and Wilhelm Pfannkuch (Kassel: Bärenreiter, 1963), 221–30 (“Dreiteilung” on 226). In 1988 Krämer translated Larsen’s *Entwicklungspartie* as “elaboration section.” Other Haydn scholars, however—such as Michelle Fillion (see below) and James Webster (note 7 below)—prefer to translate it as “expansion section.”

In our opinion, Larsen’s term “three-part division” is misleading, because it refers to the surface features of only one variant of the continuous exposition (of which we have identified several subtypes). Nevertheless, it has become a common point of reference of the current Haydn literature, as in Michelle Fillion, “Sonata Exposition Procedures in Haydn’s Keyboard Sonatas,” *Haydn Studies. Proceedings of the International Haydn Congress, Washington, D.C., 1975*, ed. Jens Peter Larsen, Howard Serwer, and James Webster (New York: W. W. Norton, 1981), 475–81. Charles Rosen also speaks of Haydn’s occasional “three-part organization” in *Sonata Forms*, rev. ed. (New York: W. W. Norton, 1988), 100–104, and provides an example with the Symphony No. 44 (“Trauer”), first movement; our analysis of the movement differs from Rosen’s. See also the citations in note 7 below.

point of the expositional space.⁷ A clear example is provided by the dogged *Fortspinnung* in the exposition of the finale of Haydn's String Quartet in B Minor, op. 33 no. 1. As discussed in Section 6 below, other familiar occurrences of this subtype are complicated through suggestive mid-expositional feints toward normative two-part expositional practices, as if the continuous exposition that ultimately results were the product of a last-moment overriding of the more normal tendency of an exposition to subdivide into two parts. Because the strength of these mid-expositional feints can vary, introductory generalizations about this subtype are difficult to make, and each case demands individual examination. For the present, we must be content merely to cite a few additional examples of the subtype: the opening movements of Haydn's Symphonies Nos. 44 ("Trauer"), 45 ("Farewell"), 96 ("Miracle"), and 103 ("Drum Roll"), as well as the first movement of Haydn's Quartet in E \flat , op. 33 no. 2 ("Joke").

In the second continuous-exposition subtype—one also often complicated through not-fully-realized gestures toward two-part practices—an early structural perfect authentic cadence (PAC) in the second key-area (typically occurring around 50 to 70% of the way through the exposition) is followed not by a true second theme but by multiple, perhaps varied or expanded restatements of the immediately preceding cadential module. These reiterations continue throughout

most (sometimes all) of the remainder of the exposition. The result is a differing sort of mid-expositional expansion section, one that keeps re-opening seemingly closed authentic cadences through varied modular repetitions. This seems to be the procedure at work, for instance, in the first movement of Mozart's String Quartet in B \flat major, K. 458 ("Hunt"). Regardless of the subtype encountered, attempts to analyze continuous expositions as if they were two-part expositions (by undertaking a fruitless search for a second theme) can lead only to a misunderstanding of their internal processes.

Both two-part and continuous exposition types share the same *tonal form*. Each modulates to the key of the dominant (or, in most minor-mode pieces, to the major mediant) and directs its linear-contrapuntal energy toward a clearly articulated tonal goal—usually the first satisfactory perfect authentic cadence (PAC) in the new key (a root position V–I cadence in which the outer voices arrive simultaneously at scale degree $\hat{1}$). Once this cadence has been achieved (and perhaps restated through a repetition of the S-theme or its cadential module), the tonal work of the exposition is over. Although more music—perhaps a good deal—may follow in order to accomplish various expressive or proportion-satisfying tasks (Koch's "ein erklärender Periode" [clarifying period] or "Anhang" [appendix], Reicha's "idées accessoires" ["accessory ideas"]), the exposition has completed its essential mission, that of providing a clearly articulated sign of closure in the new key. We thus refer to this crucial first V:PAC (or, in minor, III:PAC) as the point of *essential expositional closure* (EEC); the corresponding moment in the recapitulation (I: PAC) is the point of *essential sonata closure* (ESC). (Table 1 lists all the abbreviations used in this article.) In a two-part exposition the EEC marks the end of S; in a continuous exposition it marks the end of the central expansion section. Restated from a Schenkerian perspective: our term EEC is generally equivalent—while acknowledging occasional exceptions—to the point of completion of the *first*

⁷Cf. note 6 above. See also A. Peter Brown, *Joseph Haydn's Keyboard Music* (Bloomington: Indiana University Press, 1986), 295 ("a totally different exposition structure" for Haydn's Sonata in C minor, Hob. XVI: 20/1, in which "there is now an expansive transition that . . . dominates the entire exposition"); James Webster, *Haydn's "Farewell" Symphony and the Idea of Classical Style* (Cambridge: Cambridge University Press, 1991), 166 ("one of Haydn's special features of form: the so-called 'three-part' exposition. This centers around a long, unstable *Entwicklungspartie* or 'expansion section' in the middle, preceded by a short first group in the tonic and followed by a short, contrasting, *piano* theme and codetta in the dominant.").

Table 1. Sonata Theory: Abbreviations and Terms (with page references to definitions)

	caesura-fill (127)
C	closing zone (121)
	complex grand antecedent (139)
	continuous exposition (118)
	default (122)
	deformation (116, 128, 129, 131)
	deployment sequence (127)
EEC	essential expositional closure (119, 120, 121)
ESC	essential sonata closure (119)
FS	Fortspinnung modules (118, 133)
HC	half cadence (121)
IAC	imperfect authentic cadence (122)
MC	medial caesura (117, 121, 123)
	medial-caesura deformation (123, 124, 131)
MMS	multimodular secondary-theme zone (147)
P	primary-theme zone (121)
	P as grand antecedent (139)
PAC	perfect authentic cadence (119)
	point (zone) of conversion (133)
PMC	postmedial caesura (146)
	rhetorical form (120, 121)
S	secondary-theme zone (117, 121)
	S-deformation (145)
TMB	trimodular block (146)
TMS	trimodular secondary-theme zone (147)
	tonal form (119)
TR	transitional zone (121–22)
	two-part exposition (117)

linear fifth-progression (Zug) $\frac{\hat{5}-\hat{4}-\hat{3}-\hat{2}-\hat{1}}{\hat{1}-\hat{V}-\hat{1}}$ in the key of the dominant.⁸

⁸As is well-known, Schenker was convinced that sonata form grew out of the interruption principle, whereby (for example) an *Ursatz* of the $\frac{\hat{3}-\hat{2}-\hat{1}}{\hat{1}-\hat{V}-\hat{1}}$ variety attains the specific middleground form $\frac{\hat{5}-\hat{2} \parallel \hat{3}-\hat{2}-\hat{1}}{\hat{1}-\hat{V} \parallel \hat{1}-\hat{V}-\hat{1}}$. In Schenker's view, the first branch of the interruption structure is completed in the exposition, and its concluding $\frac{\hat{2}}{\hat{V}}$ is prolonged by the development; the recapitulation

Because they share the same tonal form, these two exposition types are distinguished by their *rhetorical form* (the manner in which a work or section is articulated in terms of modular arrangement, thematic type and topic, caesura and cadence treatment, dynamic and textural shape, and so on). This encompasses what Koch referred to as “die Anlage” (as in “die Anlage der Sinfonie,” “the plan of the symphony”), or its ordered succession of rhetorical events.⁹ This rhetorical form may be disposed according to the norms of different rhetorical plots which had accrued to the tradition of the sonata during the course of its historical development.

Although the continuous exposition is primarily associated with the works of Haydn, he also, of course, employed the two-part exposition: indeed, in Haydn's hands witty interplay between the two types became common. (Some examples are provided in Section 6 below.) Beethoven certainly favored

rebegins on $\frac{\hat{3}}{\hat{1}}$ and progresses this time to $\frac{\hat{1}}{\hat{1}}$. In the two-part exposition, $\frac{\hat{2}}{\hat{V}}$ is reached at the beginning of our second part, and it is prolonged by motion into an inner voice—that is, by the linear fifth-progression (Zug) $\frac{\hat{3}-\hat{2}-\hat{1}}{\hat{1}-\hat{V}-\hat{1}}$ in the key of the dominant. (This fifth-progression thus occurs at the second level of the middleground.) During the second part of the recapitulation, the transposition of this fifth-progression to the tonic effects the ultimate closure of the interruption structure. See Heinrich Schenker, *Free Composition*, trans. Ernst Oster (New York: Longman, 1979), 133–39.

In large part, our decision to equate the EEC with the point of the completion of the first linear fifth-progression in the key of the dominant follows Rothstein's lead (*Phrase Rhythm*, 116, and private communication). We realize that this is an enormously complex issue: for example, not all first PACs articulate the completion of a fully supported fifth-progression; moreover, under certain common conditions—thematic or cadential-modular repetition among them—the EEC-effect of the first PAC can be deferred to the second. We also realize that many standard Schenkerian analyses would identify the moment of expositional completion at the point of a later, sometimes stronger, fifth-progression. Our concern here is not so much one of expositional closure (in the broadest or fullest sense of completion) as it is one of essential expositional closure. We plan to address this topic in a separate essay.

⁹Koch, *Versuch*, Part 3, 304 (Subsection 101); see also Koch, *Introductory Essay*, trans. Baker, 199.

the two-part model but did on occasion compose a continuous exposition (as in the slow movement of the Sixth Symphony and, arguably, in the first movement of the Seventh Symphony, with its quirky cadential complications), a format that Mozart used only rarely (as in the “Hunt” Quartet, mentioned above). By the onset of the nineteenth century the continuous exposition was almost totally displaced by the two-part format. Although occasionally encountered (as in the opening movement of Mahler’s First Symphony), the continuous exposition became increasingly unavailable to composers as a normative option. Particularly because it was habitually overlooked by theorists, it became lost to historical memory—or at least to academic or analytical memory.

3. THE FOUR ZONES OF THE TWO-PART EXPOSITION

Eighteenth-century sonata expositions begin with an area of tonal stability, the *primary-theme zone* (*P*), which establishes the tonic key as a point of departure. In a two-part exposition one also finds a later, specialized musical space in a contrasting key: this is the *secondary-theme zone* (*S*), and it persists until the articulation of the EEC. The tonal function of *S* is *cadential*: its purpose is to cadence decisively in the new key.¹⁰ Upon the production of this PAC (again, one sometimes restated through thematic or cadential-modular repetition, possibly varied), the exposition may enter a *closing zone* (*C*), whose typical purpose is to reinforce the EEC,

¹⁰Cf. note 8 above. The argument in favor of considering the *first* satisfactory PAC as the conclusion of the second theme (or theme group) proper is extraordinarily complex, but it is clearly suggested in Koch, *Versuch*, Part 3, Section 101 (as translated by Baker in Koch, *Introductory Essay*, 199: “Following the cadence a clarifying period is often appended”). The same position is taken by Rothstein in *Phrase Rhythm in Tonal Music*, 116. We shall deal with the problems of determining the extent of *S*-zones—along with typical exceptions, and so on—in a separate essay.

often through a chain of cadential modules that confirm the PAC with varying degrees of strength.¹¹ The tonal function of *C* is *post-cadential*: its purpose is to solidify further the new key, often with stronger, more vigorous rhetoric, and to secure the tonal balance of the exposition as tipped decisively toward the second key area.¹²

Before the non-tonic *S* can unfold (initiating part 2 of the exposition), a musical space for it must be opened. In Allegro compositions (rapid first movements or finales, overtures, and so on) this space cannot be entered casually. Rather, *S*-space must be forcibly manufactured, through a common device of structural punctuation that we term the *medial caesura* (*MC*). A medial caesura is usually built around a strong half cadence (in the major mode either V:HC or I:HC) that has been rhythmically, harmonically, or texturally reinforced. This caesura has two functions: it marks the end of the first part of the exposition (hence our adjective “medial”), and it is simultaneously the decisive gesture that makes available the second part.

In order to accomplish this, the MC requires energy. This energy is supplied by the preceding *transitional zone* (*TR*). Within allegro movements, TR is an area of rhythmic verve

¹¹Rothstein, *Phrase Rhythm in Tonal Music*, 114–18.

¹²In our view, the name of a zone most properly signifies its *rhetorical* rather than its *tonal* function. More specifically, “*S*” refers primarily to a *thematic function* within a two-part exposition, a thematic function, we shall argue, occurring by definition *after* a clearly articulated medial caesura; the usual tonal function of this zone is linear and contrapuntal (as discussed in notes 8 and 10 above). A parallel argument might also be made for “*C*.” We believe that it is imperative to maintain this distinction between the *rhetorical* function of a zone and its *tonal* function. According to this distinction, the linear-descent tonal function of the second portion of an exposition might arguably begin before the medial caesura, but the rhetorical “*S*” proper never will. Still, speaking more casually, one might say that the (rhetorical) secondary-theme zone (*S*) is also normally superimposed over a cadential tonal function; similarly, one may also speak of the (rhetorical) closing zone as having a post-cadential tonal function.

whose purpose is to enliven the texture and drive in a series of energy-gaining rhetorical modules toward a clearly articulated medial caesura. There are many ways of invigorating the texture: a more rapid surface rhythm, an accelerated harmonic rhythm, a higher dynamic level, a more active accompaniment pattern, chromaticism, and so on. In addition, TR often features a modulation to the key of the dominant or mediant. Although some theorists have viewed modulation as the hallmark of the transitional zone, it is better regarded as only one common way of energizing the texture; many transitions do not modulate at all.¹³ A more accurate hallmark is energy-gain. In general, a modulating TR in a major-mode exposition will drive toward a V:HC medial caesura, while a non-modulating TR will drive toward a I:HC medial caesura. Because the former type of MC occurs more frequently than the latter, we refer to these two cadences as the *first-level default* (V:HC) and the *second-level default* (I:HC) for the medial caesura in major-mode compositions.¹⁴ It is

¹³One of the first (erroneously) to insist on modulation as the crucial feature of the TR (*Uebergang*) was A. B. Marx, as in his discussion of “Die Sonatenform” (the “Dritter Abschnitt” [Third Subsection] of the “Vierte Abtheilung” [Fourth Section], itself entitled “Die Sonatenform”) in the third volume of *Die Lehre von der musikalischen Komposition* (orig. publ. 1845). We have consulted the 4th (“unveränderte”) edition (Leipzig: Breitkopf & Härtel, 1868), Vol. 3. Marx’s comments on the modulatory transition are found on 223–24.

According to Marx, a I:HC medial caesura (though he did not use that term) was appropriate only to the sonatina form (204–207, the commentary on his self-composed Example 226). The larger sonata required a more expansive, modulatory *Uebergang*, concluding normally, to judge from his self-composed example (No. 257), with a V:HC. Immediately after providing the example, however, Marx (surprisingly) provided an alternative (No. 258) furnishing a V:PAC MC, which he then declared to be “eben so gut” (“just as good”) as the more normative V:HC, although it did require a different kind of *Seitensatz* to follow it. Statistically, however, the V:PAC MC occurs much less frequently in the eighteenth-century repertory.

¹⁴Robert S. Winter provides statistical evidence of the frequency of the I:HC medial caesura in Haydn, Mozart, and others in “The Bifocal Close and

also possible for the MC to be built around a perfect authentic cadence, V:PAC (a *third-level default*), a problematic event that occurs less often than either of the other two.¹⁵ We deal with the third-level default option in Section 9 below. Our main concern in this essay is with half-cadence medial caesuras in major- and minor-mode sonatas.¹⁶

The rhetoric of the two-part exposition may be schematized as:

P TR ' S / C

where the MC is symbolized by the apostrophe and the EEC (V:PAC or III:PAC) by the slash. The medial caesura and the secondary-theme zone are the defining rhetorical features of the two-part exposition. Both are lacking in the continuous exposition. This means that as a compositional or analytical construct “S” cannot exist unless a medial caesura has opened a space for it. This leads to a central axiom of Sonata Theory, one that determines with which type of exposition we are dealing: *If there is no medial caesura, there is no S*. In other words, if there is no medial caesura, we are confronting a continuous exposition, for which “S” is an inappropriate concept.

the Evolution of the Viennese Classical Style.” *Journal of the American Musicological Society* 42 (1989): 275–337. (The present authors, however, do not find the term “bifocal close” for the I:HC MC to be helpful.)

¹⁵An imperfect authentic cadence (V:IAC) may substitute for the V:PAC as a weaker third-level MC, as, for instance, in m. 77 of the first movement of Beethoven’s String Quartet in C Major, op. 59 no. 3, although that example is complicated by the immediately preceding statement (in m. 76) of what is at first taken to be a standard V:HC MC. See note 26 below.

¹⁶Because minor-mode expositions may modulate to either the mediant or (much more rarely in the late eighteenth century) the minor dominant, their defaults are somewhat more complex. Either III:HC or v:HC may serve as the first-level default, i:HC as the second-level default, and either III:PAC or v:PAC as the third-level default.

4. THE MEDIAL CAESURA: FIRST PRINCIPLES

The basic notion of a musical caesura (implying a break, gap, or cut; from the Latin *caedere*, to cut down) is elementary enough: the term may refer to any break or pause, however mild, in the texture. Our concern here is with the specially privileged, generically stylized *medial caesura: the brief, rhetorically reinforced break or gap that serves to divide an exposition into two parts, tonic and dominant* (or tonic and mediant, in most minor-key sonatas). Its effect is usually that of an emphatic pause for breath before launching the exposition's second part.

Since a medial caesura, by definition, marks the *end-point* of a TR within a two-part exposition, it must be built around a cadence: either an HC or a PAC.¹⁷ Because TR is construed as driving toward an arrival-point within a succession of phrases, that caesura-point must articulate the end of a phrase. In our judgment, the concept of the phrase is most productively understood, both historically and theoretically, as admitting only two choices for its end-point: a half cadence or an authentic cadence. It is true that a general pause (GP) gap may be produced after a deceptive cadence (DC) (as in that brilliant, sudden gasp in Haydn's "Farewell" Symphony, first movement, m. 55) or that textural breaks might occur on pre-cadential harmonies such as ii or IV. But since these harmonies do not end phrases, they cannot normally be considered medial caesuras. They cannot open S-space and divide an exposition into two parts.¹⁸

¹⁷V:IAC may sometimes—though infrequently—substitute for V:PAC, as discussed in note 15 above.

¹⁸The only exceptions to this principle (apparently rare in the late eighteenth and early nineteenth centuries) would seem to be clearly demonstrable medial caesura deformations: texturally recognizable MC-moments subjected to unusual treatment, such as one finds, for example, in the opening movements of Mozart's Symphony No. 36 in C Major, K. 425 ("Linz") and Haydn's Symphony No. 83 in G Minor ("Hen"), as well as in the second movement

Before proceeding further, it may be helpful to recall some unambiguous instances of this structural caesura. Two unequivocal examples are:

1) Haydn, Symphony No. 104 in D ("London") first movement (Example 1): medial caesura (V:HC) at m. 64, onset of S at m. 65. Notice the grand *forte* set-up, with $\sharp 4-\sharp 4-\sharp 5$ (in the key of the dominant) in the bass at mm. 56–57, and the strong laying-down of the new dominant (V/V, mm. 57–64), ending with the typical three "hammer blows" (mm. 63–64) that one often finds at the caesura-point. The Haydnesque first-level default for S ensues immediately after the GP-breath (three-quarters of a measure): a drop to *piano* and a relaunch through a restatement, or varied restatement, of the P-theme-incipit transposed to the dominant (m. 65).

2) Mozart, Symphony No. 40 in G Minor, K. 550, first movement: medial caesura (III:HC) at mm. 42–43, onset of S at m. 44. Again we find an energetic *forte* drive, locking onto the prolonged, emphatic V/III (mm. 38–42), with reiterative motivic figures. This time

of Beethoven's Symphony No. 9. In each of these pieces the rhetorical, *forte* drive to the "normal," phrase-completing MC is prematurely shattered in mid-phrase on either a predominant or a cadential $\frac{6}{4}$ chord. This dynamic collapse yields at once to an expanded passage of "caesura-fill" texture with suddenly reduced dynamics. (Caesura-fill is discussed in Section 5 below.) Most unusually, the expanded caesura-fill, seemingly suspended in the widened caesura-gap, is obliged to complete the cadential, phrase-ending role of the otherwise blocked MC. In each case the cadence thus produced is a PAC in the new key. The "Linz" example is discussed at the end of Section 9 below.

Such a procedure, exceptional in the decades around 1800, would have enormous repercussions later in the nineteenth century. In confronting a work from the mid-1800s one should be less surprised to encounter a blocked (or even suppressed) MC, followed (or replaced) by a broadly expansive "de-energizing transition" with reduced dynamics or even sounded in *diminuendo*, strikingly non-normative by earlier, eighteenth-century standards. In most cases a de-energizing transition falls inexorably to a PAC in the new key, thus unlocking the secondary-theme zone. Examples may be found in the initial movements of Schumann's Symphony No. 4 and in Brahms's Symphonies No. 2 and 3. Their historical antecedents in this practice are doubtless such works as "Linz," "Hen," and the Ninth Symphony.

the MC-figure contains two hammer blows rather than three (m. 42); the caesura (gap) lasts for one and one-quarter measures. The Mozartean first-level default for S immediately follows the MC: a drop to *piano* and a relaunch with a new, contrasting theme (m. 44).

Such familiar examples could be multiplied at length: unmistakable medial HC-caesuras immediately followed by *piano* S-themes launching a second part abound in the classical repertory. Once past these simpler examples, however, defining the term “medial caesura” turns out to be no easy matter. In the late-eighteenth-century style we encounter not only medial caesuras in various strengths, positions, and formats but also “medial-caesura deformations” for non-normative expressive purposes.

Within an expansive, Allegro composition a simple half cadence (V:HC or I:HC) is generally insufficient to create a medial caesura, at least one commensurate in strength to the length of the preceding material (P + TR).¹⁹ In most cases something more is needed by way of local enhancement. For this reason one typically finds concomitant contrapuntal, harmonic, and rhetorical gestures that call attention to the event and identify it generically as a medial caesura. Limiting ourselves for the moment to the more common situation, an MC built around a half cadence, we may say that the medial caesura is usually produced as the final moment of articulation following several measures of preparation on a prolonged *structural dominant* (V or V/V). Thus a common sequence of events is: 1) the presentation of the initial stages of TR, by and large consistently gaining in energy; 2) the attaining of the structural dominant, which is then locked onto as a literal or implied pedal; 3) the prolongation of that dominant and the rhetorical drive to the medial caesura—a

drive that usually sustains or even increases the energy accumulated up to this point; 4) the articulation of the MC proper. (Not all Allegro compositions present all four of these events: it is possible, for instance—especially in less rhetorically grand works—to sound the MC at the moment of the arrival of the structural dominant, as in the first movement of Mozart’s Piano Sonata in C Major, K. 309, m. 32.)

In order to function as a normative medial caesura, then, the half cadence that concludes TR must be additionally reinforced. More specifically:

1) The structural dominant (the arrival of which usually precedes the MC proper, often by several measures) is often approached through a chromatically altered predominant harmony that contains $\sharp\hat{4}$. This altered predominant is most often an applied chord (V/V, V⁷/V, vii^o/V, or vii^{o7}/V in root position or inversion) or an augmented sixth chord. The chromatic line $\hat{4}-\sharp\hat{4}-\hat{5}$ or $\hat{3}-\sharp\hat{4}-\hat{5}$ often appears in one of the outer voices; if an augmented-sixth chord is employed, the typical bass line is $\flat\hat{6}-\hat{5}$. The texture at this moment is characteristically vigorous, highly active; the dynamics, usually a strong *forte*, will persist or even gain in intensity in the subsequent drive to the medial caesura.

2) Once the structural dominant has been sounded, it may be rhetorically emphasized through energetic reiterations of the half cadence. The music goes through the cadence several times, reapproaching and rearticulating it, in this way helping to produce the characteristic rhetorical drive toward the MC proper.

3) The structural dominant is frequently prolonged, perhaps by neighbor motion, as part of the drive to the MC proper. This often involves alternating V with a neighboring $\hat{6}_4$, producing $\frac{5-6-5}{3-4-3}$ neighbor motion. Sometimes the neighboring $\hat{6}_4$ is supported by $\hat{1}$ in the bass, creating an apparent V–I–V alternation. (This technique should not be confused with a reiteration of the HC, No. 2 above).

4) The normative, unflagging drive in the space between the lock onto the structural dominant and the actual articulation of the MC is of cardinal expressive importance. Any attenuation of dynamics here should be viewed as counter-generic, or perhaps—especially by

¹⁹This is not necessarily the case within slow movements in sonata (or sonatina) form, especially in early examples (such as J. C. Bach Sonatas) or other small-scale (binary) movements. There a simple HC, usually followed by a brief rest, often serves to effect the MC.

the later eighteenth century—as a less common, second-level default that calls attention to itself and challenges the prevailing norm of energy-gain. Depending on the circumstances, a dynamic emptying or collapse in this space can represent a momentary crisis of energy or confidence, a sudden loss of rhetorical will, a seeming hesitation in one’s decision to enter S-space, or something similar. In such instances the S that follows the dynamically weak MC might compensate for this enervation (as is suggested also at the end of No. 7 below).

5) At the point of the MC proper one often hears several *forte* hammer blows (three is the most common number) that ostentatiously reiterate the final dominant chord. The hammer-blow effect is a common means of bringing the energy-gain of TR to a terminal peak and, simultaneously, of beginning to discharge the tension for the subsequent drop to S. The first of these hammer blows typically falls on a strong beat, often (though not invariably) on an accented measure of hypermeter. Particularly characteristic is the disposition of the hammer blows in such a way that the second (or second and third) is sounded an octave below the first. Within melodic phrase-endings, Koch referred to this formulaic octave-drop gesture (on weak beats or measures) both as a type of “Nachschlag” (a striking afterwards) and as a “Cäsur” that has been provided with an “Ueberhang” (an overhanging) or “einen weiblichen Ausgang” (a feminine ending).²⁰ Ascending octave-leaps—far more energetic—are also possible as part of this *Nachschlag* figure.

6) At the point of the MC one frequently encounters a general pause (GP) or rest in all voices. This is one of the main hallmarks of an unequivocal MC, and it usually signals the precise moment of the medial caesura. The silence of the caesura-gap is a watershed moment relinquishing the preceding drive and energy-gain. The silence articulates and represents *energy-loss*, thus initiating the subsequent, normative drop to *piano* for S. From the vantage-point of TR, however, the point is that a higher level of activity and energy has been attained: the gears have shifted, and we are now prepared to enter

the next stage of the sonata-exposition, usually at a lower dynamic level. (Note, however, that when S begins with an upbeat, as in the first movement of Mozart’s Piano Sonata in D, K. 311, m. 17—following a I:HC MC in m. 16—that upbeat might occupy the implied GP-gap. Section 5 below also discusses the common procedure of caesura-fill.)

7) Immediately following the MC proper (after the implied or actual GP-gap), one expects to find a sudden change of texture, usually combined with a precipitous drop from *forte* (at the MC point) to *piano* and the unfolding of a melody articulating the second expositional key. This abrupt dynamic/textural change strongly suggests the immediate emergence of a normative-rhetorical candidate for S-status (the launching of the second part of the exposition), an emergence that usually confirms the MC-status of the preceding HC. Particularly in large-scale compositions, this criterion is of crucial importance: the change of texture and/or dynamics functions as the standard gesture that accepts and ratifies the preceding caesura as the MC. Pointedly refusing to initiate any of the characteristic opening-types of an S-theme at this moment may signal that the preceding, proposed MC is being declined by subsequent events. (This topic is taken up in Section 7 below.)

But to this general principle of the *piano* S, far and away the standard option, one should add a word of caution. Although it cannot be maintained that the beginning of an S-theme can never be articulated at a *forte* dynamic level, in context such suddenly blurted or surging S themes are almost invariably reactive to some earlier crisis in the TR zone (especially to a crisis in the articulation of the MC, one type of which is suggested in No. 4 above). To be sure, the unusual, *forte* S may be found to great effect here and there in Mozart and Beethoven, but it seems to have been of special interest to the mature, ever-inventive Haydn, in whose works the S-ness of the *forte* theme, when it occurs, is usually identifiable through its monothematic incipit, recalling P. (It is encountered, for example, in several of the “London” Symphonies, with locally clever implications, as in the first movement of Symphony No. 99 in E \flat , m. 48.)

As an illustration of the above principles, the V:HC medial caesura in the first movement of Haydn’s Symphony No. 104 (Example 1) is reinforced by conditions 1 (approach to the

²⁰Koch, *Versuch*, Part 2, 394 (within Subsection 95 and subsequent subsections, which deal at length with the central concept of the caesura); Koch, *Introductory Essay*, transl. Baker, 23–24 (the translation used here).

Example 1. Haydn, Symphony No. 104 in D Major, first movement: mm. 54–66

The musical score consists of three staves. The first staff (measures 54-66) is for the first violin, showing a melodic line with a dynamic marking of $[f]$. The second staff (measures 58-66) is for the piano, showing a rhythmic accompaniment with a dynamic marking of p . The tempo is marked $[Allegro]$. The key signature is D major (two sharps).

dominant through V_5^6/V , with $\hat{4}-\sharp\hat{4}-\hat{5}$ in the bass, mm. 56–57), 3 (prolongation of V by $\frac{5-6-5}{3-4-3}$ neighbor motion, mm. 57–62) fortified by a constant energy-gain up to the MC, 5 (three hammer blows, mm. 63–64), 6 (general pause, m. 64), and 7 (change of texture and emergence of the new key, m. 65). The III:HC medial caesura in the first movement of Mozart's Symphony in G minor is bolstered by the same five conditions: here the dominant is approached through V^7/V (mm. 34–37), the neighboring $\frac{6}{4}$ motion is expanded to include vii^{o7}/V over a dominant pedal, and there are only two hammer blows (m. 42, including the characteristic octave-drop). Because these two medial caesuras are reinforced by the same conditions, they may be heard as roughly equivalent in strength.

In general, the more conditions involved, the more decisive the caesura. Moreover, the strength of an MC is relative to the overall proportions of the exposition. Thus, within Allegro compositions, the larger the scale of the exposition, the more conditions the HC should meet in order

to qualify as a medial caesura. By this principle, an HC-caesura reinforced only by a single condition might sound relatively strong within a small-scale exposition but relatively weak within a much larger exposition.

Another concern surrounding the identification of a medial caesura is its *temporal (proportional) appropriateness*—its precise placement within an exposition. This issue is complicated by the fact that an MC (including the possibility of the third-level default, V:PAC) could occur anywhere from about 15% to 70% of the way through an exposition. To be sure, this is a broad expanse of expositional space, even though most cases fall before the halfway point. Our research suggests that the deployment of the I:HC MC is flexible, occurring typically within the 15–45% range: what is noteworthy here is the early availability of the I:HC MC. Beyond the 45% point—and especially in grand-scale works, such as symphonies, often earlier than this—the I:HC MC seems to have been considered either eclipsed altogether or increasingly and rapidly left behind as a practical option. Thus the

second-level-default I:HC is the first temporally available MC-deployment option within an exposition. (Haydn's expositions, for example, sometimes make an early structural-dominant or even MC-like feint toward this I:HC option only to renounce it or pass it by for a later V:HC or V:PAC MC. In such cases the witty effect, found in many of the "Paris" and "London" Symphonies, can be that of demonstrating the compositional options that he is choosing not to deploy.)

The normally available range for the more common, first-level-default V:HC MC overlaps broadly with that of the I:HC MC but in general occurs slightly later: when selected, the V:HC MC option is typically placed from about 25 to 50% (more rarely 60%) of the way through the exposition.²¹ (In Example 1, from Haydn's Symphony No. 104, the V:HC MC concludes at the 45% point in m. 64, the 48th measure of a

²¹From time to time Haydn's obsessive drive for unpredictable or non-normative originality produces extraordinary exceptions to this principle. What appears to be the V:HC MC in the first movement of Symphony No. 82 in C Major ("Bear"), for instance, occurs in m. 69, 68% of the way through the exposition. In part, this occurs as a result of Haydn's earlier dalliance with the I:HC option (notice the bassline's lock on V/I, mm. 33–39, gradually loosened in subsequent measures), his slowly-unfolding rejection of that option, and his witty composing-out (with bearish growls in mm. 51–52?) of the prolonged difficulty of finding the new V/V structural dominant, which is finally attained only in m. 59 (as one of the consequences of an unusual, preceding drop to *piano* [m. 57] and marked also with a thematic module in the upper voice that anticipates the new-theme-to-come at m. 70). Larsen, "Sonata Form Problems," 274, takes the following *piano* theme, m. 70, to be an archetypal example of the opening of the third part of a Haydn-esque "three-part division of the exposition" (what we call a "continuous exposition," as discussed in note 6 above). In this case we disagree. Based both on the peculiar rhetorical narrative produced in TR and on the acceptably S-like rhetoric of the *piano* theme at m. 70, anticipated in the preceding drive to the MC, we believe it preferable to understand that theme as an extraordinarily late S. (The first movement of Symphony No. 97 in C, which Larsen also regards as an archetypal example of expositional *Dreiteilung*, presents remarkably similar, though ultimately more problematic issues. There the S-status of the analogous theme is not quite so convincing.)

107-measure exposition; in Mozart's Symphony No. 40, the III:HC MC—the minor-mode equivalent of the V:HC option in major—occurs at the 42–43% point.) Correspondingly, the third-level-default V:PAC MC option is usually located in the 50–70% (very rarely, 75%) range. This is the last available deployment option, and it is sometimes encountered as a recovery from a failed attempt at producing a V:HC MC. Any strong caesura falling outside these boundaries is either an exceptional MC (in which case a cogent argument on its behalf would have to be offered) or, more often, no MC at all.

More important than the precise percentage numbers, which admit of exceptions and may be adjusted through subsequent research, is the overriding principle of the normative *deployment sequence* of potential structural dominants and/or MCs: the initially available I:HC soon overlaps with and eventually gives way to the V:HC option; if the V:HC option is not selected, the last chance to produce a two-part exposition resides with the possibility of articulating an appropriately placed V:PAC MC. Any relatively late V:PAC MC brings with it certain structural complications and potential ambiguities. We defer a discussion of these until Section 9 below.

5. BRIDGING THE GAP: CAESURA-FILL AND CAESURA DEFORMATION

Condition 6 above is the general pause (GP). This pause in all voices is usually short, lasting at most a measure or two or, more often, less than a measure. But sometimes that brief gap—whose very frequency in the eighteenth-century style gave rise to the term *caesura* in the first place—is filled by bridging material in one or more voices. This bridging material might be a sustained dominant pitch or chord in the upper parts, the initiation of a new accompaniment pattern, a short melodic descent leading to the initial tonic of S, and so on. Our term for this material is *caesura-fill*.

An elementary instance of caesura-fill occurs in the first movement of Haydn's Symphony No. 100 in G ("Military"). Here the MC (V:HC) is reached at the downbeat of m. 73.²² Most of the orchestra stops playing at this moment, producing the usual caesura-gap. Above, however, a flute sustains a two-measure a^2 (mm. 73–74, marked with a trill in some editions) bridging the gap, and two oboes enter as harmonic reinforcers of the V^7 in m. 74. S itself begins at m. 75, in the same flute and oboes: typically for Haydn, this second launch with S features a transposed variant of the P-theme. Mm. 73–74 imply silence but in fact are filled with sound. In this and similar cases we would not consider caesura-fill to be a deformation (an overriding of a normative option). On the contrary, it is a common gesture, part of normal practice—merely one way of articulating an implied GP or guiding the ear through a structural gap. In nearly all instances, the fill articulates or stands for the energy-loss of the GP as it leads to a gentler S-theme, explicitly or implicitly marked with a *piano* dynamic.

Other types of caesura-fill can lead to more complex issues. In one characteristic type the caesura-fill leads from the V:HC MC down to the tonic pitch of the newly established V, as though the fill's task were to lay down the tonal platform on which S will make its appearance. Because this procedure often involves an outer-voice melodic descent from $\hat{5}$ to $\hat{1}$ in the new key, we refer to this gesture as a *caesura-fill of the $\hat{5}$ – $\hat{1}$ descent type*. At times it might suggest something cadential (the V:HC is being led to an implied authentic cadence in V), but in most instances such an event is better considered a secondary, linear move that directs our attention

²²The structural V/V is articulated at the downbeat of m. 62, after which it is prolonged. During this prolongation the seventh is added (entering first in mm. 64–65, though most prominently in mm. 69–73), suggesting a V^{8-7} figure. The seventh ($\hat{4}$) resolves to an inner-voice $\hat{3}$ at the onset of S (m. 75). This addition of the seventh during the drive to the MC proper is not uncommon.

from the harmonic interruption on V (the caesura) to the restart on the new tonic that follows.

A simple case may be found in the finale of Mozart's Piano Sonata in B \flat , K. 281. (Labelled a "rondeau," this movement is a sonata-rondo mixture with a normal exposition through the S-point.) Here we find a relatively light V:HC medial caesura in m. 27; it is complemented with caesura-fill in the left hand, outlining a $\hat{5}$ – $\hat{6}$ – $\hat{5}$ – $\hat{4}$ – $\hat{3}$ – $\hat{2}$ – $\hat{1}$ linear motion in F major, mm. 27–28; the onset of S occurs at the downbeat of m. 28.²³ The $\hat{5}$ – $\hat{1}$ descent or a relevant portion thereof ($\hat{5}$ – $\hat{3}$, for example) may also be found in an upper voice, in which case it will effect a melodic link between the end of TR and the beginning of S.²⁴

In K. 281 the $\hat{5}$ – $\hat{1}$ linear fill occurs in only one voice; the other voices shut down until the S-launch. In such cases the GP and the harmonic interruption can be clearly felt; there is no compelling sense that the caesura-dominant has resolved to the new tonic, producing a PAC. Nor is the situation much complicated by such standard occurrences as that in the first

²³We follow the reading in the *Neue Mozart-Ausgabe*, "Klaviersonaten," ed. Wolfgang Plath and Wolfgang Rehm, vol. 1 (Kassel: Bärenreiter, 1986). As Plath and Rehm mention, however (35)—and question as an error—the autograph score shows the motion a^1 to $b\flat^1$ instead of a^1 to g^1 in the middle voice of m. 27. If the autograph reading is accepted, as it is in some other editions, the passage would be a clear instance of an HC on V^7 (V^{6-7}_{+3}) rather than on V. In this case the seventh would enter immediately (cf. note 22 above).

²⁴From a Schenkerian perspective, a $\hat{5}$ – $\hat{1}$ fill in the upper voice differs qualitatively from the same event in the bass voice. Because most S-themes exhibit a $\hat{5}$ – $\hat{1}$ or $\hat{3}$ – $\hat{1}$ descent culminating in the EEC, a theme that begins on $\hat{1}$ is understood as starting on an inner voice and pursuing an initial ascent (*Anstieg*) to the primary tone $\hat{5}$ or $\hat{3}$. Thus a $\hat{5}$ – $\hat{1}$ caesura-fill in the melody represents motion into an inner voice, while a $\hat{5}$ – $\hat{1}$ fill in the bass generally leads to the real bass tone (the root of the new tonic). In Section 9 below we suggest that the less common V:PAC (or III:PAC) medial caesura may often be a highly developed instance of the $\hat{5}$ – $\hat{1}$ HC caesura-fill. If this is true, the interpretation as a motion into an inner voice would help explain why the PAC of this caesura does not effect the EEC.

movement of Beethoven's Piano Trio in E \flat , op. 1 no. 1, mm. 31–32. Here the piano and violin shut down after the MC (V:HC) at the downbeat of m. 31; the cello supplies a two-measure $\hat{5}$ – $\hat{1}$ descent, although the caesura-fill is texturally enriched by the re-entry of the violin in m. 32, preceding the beginning of S at m. 33.

The situation becomes more arresting, however, when the principle of linear ($\hat{5}$ – $\hat{1}$) caesura-fill undergoes a deformation (a pointedly non-normative treatment for expressive effect). In the first movement of Beethoven's Piano Trio in G, op. 1 no. 2 (Example 2), we may easily recognize the operative principle at the MC–point. A clear approach is made to what we expect to be a normative triple hammer-blow V:HC MC at mm. 97–98. At this juncture the violin and cello drop out for the remainder of the measure, while the right hand of the piano part traces out a melodic fill from g^2 down to d^1 . More important, the usual caesura-fill energy-loss is absent here. On the contrary, the fill, continuing in aggressive triplet-sixteenth-notes, insists on retaining the full measure of gained energy and plunges precipitously to the new D-major tonic, now reinforced by the strings (m. 99), before S itself emerges, *piano*, at the upbeat to m. 100. The composer has wrenched a normal MC, V:HC (first-level default), into a strong V:PAC (third-level default) by brute force.²⁵ Beethoven's specific technique involves a textural transformation (deformation) of the $\hat{5}$ – $\hat{1}$ linear caesura-fill principle. From one perspective,

²⁵The structural tone at the end of TR is a^2 ($\hat{2}$ in the tonic key of G), gained at m. 89 in the piano. Beethoven then executes a V^{8-7} motion, during which a^2 moves to g^2 . This g^2 (heard most prominently in the piano part of mm. 97–98) resolves to the violin's $f\sharp^2$ on the second beat of m. 99 (a quarter-note delay of its conceptual resolution on the first beat)—that is, at the beginning of the pickup to S, after the V:PAC has been articulated. The resultant a^2 – g^2 – $f\sharp^2$ line thus represents motion into an inner voice ($\hat{5}$ – $\hat{4}$ – $\hat{3}$ in the dominant key of D); S begins on this inner voice $\hat{3}$ ($f\sharp^2$) and quickly ascends to the *Kopftón* $\hat{5}$ (a^2) in order to begin its $\hat{5}$ – $\hat{1}$ descent. Thus, at a deeper level the piano's breathtaking cascade from g^2 to d^1 may be understood as part of an elaboration of a melodic $\hat{5}$ – $\hat{4}$ – $\hat{3}$ caesura-fill.

the V:HC still sticks as the real conceptual MC, followed by unusually strong fill. From another perspective, the force driving to the V:PAC demands consideration as a third-level-default MC obliterating the effect created a measure earlier. In this second interpretation, Beethoven would be understood to transplant—as if by heavy lifting—an already-established caesura from one point to another (“Not *there* . . . but *here!*”).²⁶ Both the vehemence with which this event is treated and the forceful manner in which characteristic exposition-norms are overridden are typically Beethovenian. Sonata Theory helps us to articulate what Beethoven appears to be doing at this moment.

Another enhanced MC-situation occurs when the caesura-fill involves all (or most) of the voices. One example may be found in the first movement of Mozart's Symphony No. 39 in E \flat , K. 543. Here one encounters a clear MC (V:HC) with octave-drop *Nachschlag* in m. 90. Mozart suggests in this case, however, that the MC-articulation is insufficient to stop the juggernaut, triple-time momentum. Directly with the sounding of the MC there ensues an expanded caesura-fill, featuring $\hat{5}$ – $\hat{4}$ – $\hat{3}$ – $\hat{2}$ – $\hat{1}$ linear motion in the strings, presumably still *forte* (mm. 91–97), that arrives on the new tonic B \flat (*sfp*) with the effect both of a pseudo-cadence and, perhaps, of a finally-exhausted gasp (m. 97). The ensuing S (apparently with its metrical head spinning from the effects of what has preceded it) begins at m. 98, *piano*, in the key of the dominant.

This last example illustrates a situation in which the normally-brief caesura-gap is not only filled but stretched and prolonged, delaying the onset of S, in this case doubtless for

²⁶Beethoven replicated this effect in the first movement of the String Quartet in C Major, op. 59 no. 3, mm. 76–77. In this case, however, the caesura-fill is produced by the two upper strings in aggressive parallel thirds, and the cadence produced by force is V:IAC, not V:PAC. A *piano* S ensues on beat 2 of m. 77.

Example 2. Beethoven, Piano Trio in G, op. 1 no. 2, first movement: mm. 93–103

93 [Allegro vivace]

98

sf *f* *p* *f* *f* *p* *f* *f*

a super-energized, juggernaut effect.²⁷ For sonatas in the 1780s such an expansive caesura-fill is probably best considered a *medial-caesura deformation*, that is, a strikingly unusual or strained procedure relying on our knowledge of the typical limits of the norm. It is not the caesura-fill principle itself that is generically deformational here, but rather the pulling-apart of the caesura-gap. Caesura-fill—normally a mild bridging effect—and caesura deformation are different things, even though composers seeking the more powerful effect of a caesura deformation often use caesura-fill to accomplish their expressive ends.

Although in Austria at the time of Mozart's Symphony No. 39 such an *expanded caesura-fill* may have been deformational, this situation was subject to change in the hands of any later composers who determined to make it a common option within their own personalized styles. This is the case with Rossini a quarter of a century later. One of the standard features of Rossini's schematic, gridlike approach to opera overtures was the crystal-clear articulation—even the witty overarticulation—of a standard, first-level-default MC (HC in the key of S-to-come), followed immediately by an ex-

panded, separately thematized caesura-fill giving the effect of a momentary suspension of time or a quiet, guided passage through the caesura-void.²⁸ Rossini's Overture to *La gazza ladra* (1817) in E minor, shown in Example 3, may be considered paradigmatic. After multiple stutterings (whose extravagant reiterations also produce a palpable loss of dynamic energy), a comically emphatic MC (III:HC) is produced at m. 159. It is followed by a measure of silence (m. 160) and eleven suspended measures of a hushed, operatically expectant caesura-fill of the $\hat{5}-\hat{1}$ type (mm. 161–71). For the moment, as is common in Rossini overtures, sonata time seems to stop; it snaps back into action only with the onset of a normative S at mm. 171 and 172. Within Rossini's personalized customization of the style, this was no caesura deformation; it was the norm.

This does not mean that all expanded caesura-fills were normative by the 1810s. Weber's Overture to *Der Freischütz* (1821) presents us with an exceptional set of expansions, ambiguities, and deformations at the apparent MC-point. Toward the end of TR, the C-minor overture modulates abruptly to E \flat major at m. 87 and throws itself impetuously into a *fortissimo* PAC in that key at m. 91. This leads to harmonic stasis and (with the exception of the unsettling E \flat chords in the horns, mm. 93–96) an immediate reduction in sonority: *piano*, tremolo strings sustaining the newly-produced E \flat tonality. On the face of it, m. 91 would appear to be a masked articulation of an impulsively-gained, structurally premature third-level-default MC (III:PAC).²⁹ Yet a considerable stretch of subsequent music, mm. 91–122—surely not yet the rhetorical S idea, despite its E \flat tonality—precedes the onset of the obvious S—proper, the major-mode

²⁷Not surprisingly, Beethoven was especially fond of juggernaut caesura-fill—an expressive type that need neither be expanded to multiple voices nor stretch the caesura-gap. In the first movement of the Second Symphony, for example, one finds an aggressive TR leading to a normative initiation of an MC (V:HC) at m. 71. One might have expected here the stereotypical three hammer blows on the dominant (three quarter notes followed by quarter rests, plus a half rest in m. 72). Instead, impetuously, the ongoing motion and dynamic level charge directly *through* the caesura-point (mm. 71–72) with a variant of the $\hat{5}-\hat{1}$ linear descent. Here Beethoven suggests that the impetuosity and energy that have been built up will not be easily containable. There follows a normative drop to *piano* and change of texture for the march-like S (m. 73). The energy-level has been built to such a point, however, that this S virtually erupts out of its *piano* casing with the *ff* consequent phrase at m. 77. Once one has obtained an awareness of MC norms, one is likely to conclude that the very strangeness of the caesura and S-behavior is just what we are supposed to be perceiving at this point.

²⁸Philip Gossett discusses the Rossini formula (though without mentioning the caesura-fill procedure) in "The Overtures of Rossini," *19th-Century Music* 3 (1979–80): 3–31.

²⁹See note 16 above.

Example 3. Rossini, Overture to *La gazza ladra*: mm. 149–73

[Allegro]

149 [ff] sf sf ff

152 pp

157 [pp] sotto voce dolce

167 [pp] [pp] [pp] [pp]

“Agathe” theme (m. 123). Nor does the music from m. 91 to m. 122 seem like a continuation of TR. Above all, it is in the new key throughout ($E\flat$) and is largely produced at a reduced dynamic level (expressing, perhaps, a prolonged suspension of lost energy or an ongoing energy-loss): this is the opposite of the defining feature of the TR (a generally consistent sustaining or increase of energy up to the MC).

As such, if one is to hear mm. 91–122 as in dialogue with the rhetorical norms established in a wide variety of preceding works (the central premise of Sonata Theory), it may be best to consider them as a monumentally expanded caesura-fill following an unusual III:PAC MC at m. 91. This caesura-fill also accomplishes the task of undoing the rash III:PAC and reactivating a III:HC at m. 122. (This is a gesture of recovery but not a typical medial caesura by any established norm). Since this is a programmatic overture with themes referring to characters and events in the opera, the expressive point at hand seems evident. Here Weber has just had the tormented Max (represented by the C-minor P-theme and storm within TR) jump rashly at the first $E\flat$ -major MC opportunity, prematurely producing a III:PAC at m. 91. Now Weber pulls the caesura-gap wide open—stops sonata time—and fills most of the caesura-void with music representing the trembling Max peering fearfully into the blackness of the Wolf’s Glen abyss. Similarly, the eventual undoing of the potential closure of the premature III:PAC MC coincides with the approach of Agathe as S—for Max, the principle of unmerited salvation. It is a magnificently poetic moment: here structural deformation and expressive (even pictorial) purposes are brilliantly merged.

6. TWO-PART OR CONTINUOUS? THE BAIT-AND-SWITCH TACTIC

When first confronting an eighteenth- or early-nineteenth-century exposition, our most reasonable expectation would be that we are about to encounter the far more common

type—the two-part exposition with MC and subsequent S. When we are presented instead with a continuous exposition of the expansion-section subtype, there is usually a moment of psychological conversion (provided that we are aware of our interpretive options)—a personal understanding at some mid-expositional point that the more standard, two-part form is not going to be realized. We believe that this expectation may have been shared by the competent listener in the decades surrounding 1800 and that Haydn, in particular, often made the process of conversion into a central feature of his pieces with continuous expositions. The mechanism through which this conversion is suggested cannot be investigated without understanding the norms surrounding medial caesuras, for in most cases of the continuous exposition potential MCs are first suggested, then abandoned. Haydn, in particular, frequently shows us the process of psychological conversion from one exposition type to the other. Demonstrating this process rhetorically is often what the exposition seems to be about.

Thus as we (as listeners) move through most later-eighteenth-century continuous expositions, what we at first suppose is an ongoing TR (on its way to an MC) continues past the last-possible S-point, or what we might designate as the *point of conversion*. (This may also be described as a brief *zone or process of conversion*.) Sensing that TR has crossed through this conceptual point or zone forces our reassessment of what is occurring generically. We come to realize that we are dealing instead with an expansion section, probably one grounded in a succession of *Fortspinnung* modules (FS) that will occupy the large center-portion of the exposition, now understood as a continuous, not a two-part, exposition. We can schematize our experience of this large, central section as *TR-FS*: what begins (we think) as TR shifts conceptually to the FS modules characteristic of the continuous exposition (an FS that will drive, without an S, toward the EEC); the hyphen represents the process of conversion.

At the basis of this notion is the assumption that a listener adequate to the basic demands of the piece actually does sense such a process (or point) of conversion. Sensing it depends both on a solid experience of the style—having a large inventory of normative exemplars at hand—and on grasping the proportions that a composer seems to promise at or near a piece's outset. For the listener, one important function of the opening ideas of each exposition is to help predict the rhetorical scale that will follow: some sonatas are brief, while others (with vaster P and TR zones) are monumentalized.

Once we have attained the ability to project the proportion-to-come, there does occur a point during the course of the presumed TR where we begin to expect a locking onto a structural dominant and a subsequent drive to a medial caesura. The TR–FS can pass through the zone of conversion in a number of ways. We may imagine the manifold possibilities as arranged on a sliding-scale representing the various degrees to which we sense that a potential MC has been suggested. For heuristic purposes we might identify three situations within this sliding-scale (in which each case pushes the sense of an MC toward a clearer articulation): 1) the TR–FS can move past the last-possible S-point with no caesura signals whatever; 2) it can reach and perhaps prolong the structural dominant— even initiate a clear, generic drive to the MC—but fail to crystallize out a medial caesura; or 3) it can actually articulate a seeming (or potential) MC and perhaps even enter a process of caesura-fill but then both decline to furnish an immediate, subsequent S and refuse to drive toward a more acceptable MC in the ensuing measures. Strictly considered, this last case, which is sometimes difficult to distinguish from extreme examples of the second, belongs to the category of medial caesura declined (Section 7 below), but the psychology of its production is perhaps best understood in its relation to the first two cases. What is needed at this point is a closer look at each possibility.

1) FS may move past the S–point without our noticing it. In other words, we eventually come to realize that we are beyond the S–point. By all reasonable standards, it is now too late for an S–theme, although we did not register our having passed by its potential moment: we heard neither a medial caesura nor any compelling generic signals of an approach to one. To be sure, such pure instances of the continuous exposition are rare among celebrated works of the later-eighteenth-century composers—the *Presto* finale of Haydn's Quartet in B Minor, op. 33 no. 1 is a *locus classicus*—but they do appear in pieces from the earlier part of the century. Elementary examples may be found in some of the Sammartini symphonies from around the early 1740s and in several of the first movements of C. P. E. Bach's keyboard sonatas from the same time, such as the "Prussian" (1740–43) and "Württemberg" (1742–44) Sonatas.³⁰

2) The composer may create the expectation of an imminent MC only to veer away from it for more *Fortspinnung* or other elaboration. How close we get to the implied caesura-point varies from case to case. The MC–point proper, of course, results from the laying-down of the struc-

³⁰In the first movement of Sammartini's Symphony "No. 3" in D Major (J-C 15, before ca. 1742) the first half of the binary (proto-sonata) structure may be construed as: P (mm. 1–8); a short-winded FS (mm. 9–20) that never suggests anything caesura-like but does lead to the EEC (V:PAC) at m. 19; a brief, cadential close (C, mm. 20–28). The score is available in *The Symphonies of G. B. Sammartini: Vol. 1: The Early Symphonies*, ed. Bathia Churgin (Cambridge: Harvard University Press, 1968), 76–77.

In C. P. E. Bach one often finds a format similar to that mentioned above: an initial P–gesture; a modulatory FS (typically sequential—and rarely very long) that proceeds to a PAC (the EEC); and a (brief) "appendix" theme (C) at the end to solidify the new key. Because C. P. E. Bach's textures so often feature breaks and discontinuities, the caesura situation is sometimes difficult to assess. For a general discussion of C. P. E. Bach and the frequent inappropriateness of the concept of the second theme, see David Schulenberg, *The Instrumental Music of Carl Philipp Emanuel Bach* (Ann Arbor: UMI Research Press, 1984), e.g., 100–105; and William S. Newman, *The Sonata in the Classic Era*, 2nd ed., 420–21.

tural dominant, the harmony that could potentially articulate a I:HC, V:HC, or III:HC medial caesura. The structural dominant may be touched lightly and immediately rejected (as if hot) with a new burst of *Fortspinnung* that overrides (or writes over) the normal tendency of the exposition to divide into two parts at this mid-expositional point. In other cases one locks onto the structural dominant and approaches the production of an MC—begins to fall into one—then draws away from it before that MC turns into a reality.

An example of the latter situation is provided in the first movement of Haydn's Quartet in E \flat , op. 33 no. 2 ("Joke"; Example 4). Here TR sets out in the tonic in m. 13 and moves almost immediately to V/V on the third beat of m. 14. This newly locked structural dominant now underpins a generic drive to what we presume will be a standard V:HC MC, a drive beginning in earnest with the reiterated figures in m. 15. The reiterations and hypermetrical implications clearly suggest the production of a normative medial caesura in m. 19: it would be easy to imagine a differing m. 19 that consists (assuming the most generic of choices) of three hammer-blow F-major chords (V of B \flat), followed by a rest, a drop to *piano*, and (since this is Haydn) a monothematic S theme—or perhaps a contrasting one—beginning with the upbeat to m. 20. But instead, at the last moment, in m. 19, Haydn slips out of the caesura-loop by sustaining the first violin and cello, unsettling the immediate dominant (thus inaugurating a new harmonic progression at the precise moment when we had expected everything to stop), and gliding forward into a re-invigorated melodic figure in the outer voices. This new figure (obviously grounded in much that has preceded it) is immediately imitated in the second violin, and then, in m. 20, in the viola. In short, a renewed thematic idea emerges and pushes through the MC-moment (writes over it), cancelling the local MC-implications with a new burst of *Fortspinnung*. Mm. 19–20 represent the point (or zone) of conversion, the point at which a two-part exposition is renounced, and the

Fortspinnung continues by merging smoothly into a cadential module beginning on the new tonic in m. 21 and expanding outward until the EEC is attained on the third beat of m. 28. The exposition itself ends four measures later, in m. 32. (The weak V:PAC at m. 21 should not be considered the EEC: m. 21 is a direct and relatively uninterrupted continuation of the figuration of the preceding measures. This PAC is probably better understood not as concluding anything but as marking the tonic-chord onset of a thematically profiled cadential module, a common feature of the conclusion of Haydn's expansion sections.)

As a whole, this passage from op. 33 no. 2 illustrates the procedure that we call the *bait-and-switch* tactic: Haydn baits us into anticipating an imminent medial caesura, the hallmark of the two-part exposition, then swerves away from the caesura-point and switches to a continuous exposition of the expansion-section subtype—all for the sake, one supposes, of high generic play and the splendid exhilaration found in sophisticated musical humor.

3) In extreme cases of the bait-and-switch tactic we find the MC fully articulated before the plug is pulled on the two-part exposition. Such a situation occurs in the first movement of Haydn's Symphony No. 96 in D ("Miracle"; Example 5). Setting aside the delicious complications that bring us to the V:HC MC point (including a typically Haydnesque attempt to re-open the I:HC MC possibility in mm. 48–51, aborted in m. 52, perhaps because the I:HC option had already been used up earlier in m. 31), we may note that mm. 54–55 drive to the new structural dominant, V/V, which is attained in m. 56. This leads to the manufacture of a nearly immediate V:HC MC with upward *Nachschlag* on the first beat of m. 57, followed at once by an eighth-rest. (As a result of the earlier complications, this is an exceptionally late first-level-default MC, occurring, as we eventually learn, some 61% of the way through the exposition, if we consider the expositional space as continuing through m. 83.) The upbeat

Example 4. Haydn, Quartet in E \flat , op. 33 no. 2 ("Joke"): mm. 13–28

[Allegro moderato]

13

18

23

26

tenuto

cresc.

f

p

f

tr

Example 5. Haydn, Symphony No. 96 in D ("Miracle"), first movement: mm. 54–71

to m. 58 in the strings, with its characteristic energy-loss drop to *piano*, begins a recognizable expanded caesura-fill in octaves. Its upward motion, however, is non-normative, gaining rather than losing registral energy. Consequently, the caesura-fill is made to overshoot its tonic-pitch goal in m. 60, then to draw itself up questioningly on $\hat{4}$ of A major (m. 61), and finally to abandon the fill function altogether with the incongruous intercutting of a *sforzando* G^7-C progression (momentarily calling our attention to $\sharp III$ of the anticipated A major) in mm. 62–63. The top voice of this C-major chord recaptures the e^2 of the MC *Nachschlag* (m. 57), whereupon a descending fifth progression (from $\hat{5}$ in m. 63 through $\hat{4}-\hat{3}-\hat{2}$ in mm. 64–65 to $\hat{1}$ in m. 71) leads to a V:PAC in m. 71. The cadential $\hat{4}$ of m. 67 recovers the dominant of A major, now understood as having been prolonged from m. 56. The V:PAC of m. 71 is no late medial caesura. Appearing some

82% of the way through the exposition and eliding with a clearly codetta-like C theme, it serves unambiguously as the EEC. Here the witty zone of conversion from a two-part to a continuous exposition is best heard as occurring in mm. 61–63: what began as caesura-fill is converted into a structural linear descent. Notwithstanding the pointed MC in m. 57, by m. 63 it seems clear that the potential two-part exposition has been discarded.

Such bait-and-switch procedures as we find in op. 33 no. 2 and Symphony No. 96 are typical of Haydn's continuous expositions, most of which are grounded in gestures toward two-part expositions that are abandoned to pursue other structural paths. What differs from case to case are the unfailingly engaging details and the degree toward which the jettisoned two-part proclivities remain perceptible through the continuous musical surface. Haydn's inventiveness along

these lines never ceases to astonish. The general psychology at work—seeming to promise one thing but delivering another—is at the core of his imagination as a master composer.

7. MEDIAL CAESURA DECLINED

Section 5 considers the possibility of a *caesura deformation*, the situation in which a medial caesura is altered in some non-normative way. In none of the examples provided in that section was either the MC itself or the resultant two-part exposition called into question. As we have seen with the “Miracle” Symphony at the end of Section 6, however, it is possible for a composer to create the impression that the music *following* an apparent MC (or MC candidate) conceptually undoes that caesura by refusing to accept its implied consequences. This would be a *retrospective cancellation*: a medial caesura has been *proposed* but the subsequent music has *declined* to accept it—has declined, that is, to initiate the second part of a two-part exposition, preferring instead to continue unfolding under the structural categories of the first part (P + TR). In general, we consider a medial caesura accepted if what follows it is a satisfactory S-theme. When what follows it is not, there arises the situation that we call *medial caesura declined*.

Before addressing the central question in all instances of medial caesura declined—how may we decide what constitutes a satisfactory S?—some preliminary remarks might be helpful. Most often, the declined MC occurs early in the exposition and is articulated as a I:HC, the harmonically weaker, second-level default. Nor is this surprising: although the I:HC MC furnishes less harmonically decisive structural punctuation, it is also the earliest available MC candidate within the normative expositional deployment sequence of structural dominants and/or MCs (as outlined in Section 4 above). To decline a I:HC MC candidate need not initiate

a transitional crisis, since the next-available HC MC option, the stronger V:HC, remains available shortly down the road. Still, one cannot confine the category of medial caesura declined only to the I:HC: declining a V:HC MC candidate can occur, although it is less common. To reject an early I:HC MC, however, suggests that had it been accepted (and had it proceeded directly into S in the new key) the proportions of the remaining exposition-to-follow (part 2) would have been correspondingly brief. Regardless of the default-level of the proposed caesura, the expressive purpose of medial caesura declined is normally to show the compositional decision to spring into a proportionally larger frame—the decision to manufacture a grander, perhaps monumentalized exposition (and hence movement as a whole). Medial caesura declined suggests the musical equivalent of the statements, “No! Let’s produce something larger!” or (when a I:HC is at stake) “No! The weaker caesura-option won’t do! We need something stronger!”

Because the first MC candidate’s invitation to enter directly into S-space has been rejected, the subsequent music can proceed as if there had been no structural caesura. The usual strategy is to reinvigorate TR-activity (fairly soon) to produce a real MC before too long—almost always a stronger one, V:HC. It frequently happens that such an MC can readily be found and that the now-expanded exposition will follow it with an appropriate S-theme. But such continued TR activity cannot go on at length. Expository time is running out, and the music will soon approach the proportional point at which it must renounce the possibility of providing a medial caesura (and hence a two-part exposition) at all. Thus a medial caesura declined always raises the possibility of conversion from a two-part to a continuous exposition. Each case must be considered individually.

Given a proposed MC, what kinds of evidence point in the direction of medial caesura declined? What would encourage us not to understand the material that follows it as the obvious

S? This should happen when the relevant passage exemplifies at least one of the following three situations:

Situation 1: *Following a proposed I:HC MC, the music refuses to leave the tonic key (perhaps even restating P-material).* This may seem a self-evident point, but it conceals latent complexities, not all of which will be elaborated here. In brief, Situation 1 can occur in at least three degrees of architectural strength, each of which harbors multiple implications within Sonata Theory for an adequate understanding of the initial modules of the P + TR block. The weakest instances encompass I:HC quasi-caesura gestures that may not be genuine MC candidates at all, but only the concluding elements of the common exposition-launching strategy that we call *P as grand antecedent*. In its simplest manifestations, of course, an antecedent is conceived as a single, brief phrase. Yet it is not uncommon for expositions to begin with P-themes that feature a lengthy, typically *multimodular* antecedent idea of more than 12–16 measures containing several sub-phrases or subparts linked together, often arranged in some variant of sentence-form (aa'b). (It is also possible that the sentence's presentation phase, aa', might itself unfold as a small-scale period. Such a situation produces a hierarchical nesting of different thematic shapes within differing levels of architecture: the small period comes to be reconceptualized as the presentation phase, aa', of a larger sentence, which in turn serves as the grand antecedent of an implied grand period.) Such a large, multimodular antecedent frequently drives to its end-point, a I:HC, with bold, energy-gaining, rhetorical flourishes—sometimes even MC-like flourishes. The very breadth of such a P suggests the striving for monumental proportions. In most cases, the grand antecedent will lead to the onset of a *parallel grand consequent* (marked by a restatement of the incipit of P in the tonic) that soon dissolves into more normative TR rhetoric.

Paradigmatic examples of the P-as-grand-antecedent strategy include the first movements of Mozart's Symphonies

No. 40 in G Minor, K. 550 and No. 41 in C Major, K. 551 ("Jupiter"). In No. 40 the dominant is reached in m. 16 and expanded with flourishes for four measures; in No. 41 the dominant is reached in m. 19 and elaborated for four measures up to a *fermata* HC close in m. 23. Both lead directly to a dissolving parallel grand consequent, construed in Sonata Theory as the onset of TR proper (TR of the dissolving-consequent type). The opening of Haydn's Symphonies No. 82 in C Major ("Bear") and 83 in G Minor ("Hen") present similar situations: the I:HC quasi-MC effect occurs in m. 20 in the former, in m. 16 in the latter; again, each is followed by a reiteration of the beginning of P in the tonic. One might cite also the first movements of Beethoven's "Waldstein" and "Appassionata" Sonatas—examples of grand antecedents are not difficult to find. Probably in none of these cases would one understand the HC-conclusion of the grand antecedent as a genuinely proposed I:HC MC inviting one to enter the secondary-theme zone. If this is so, the category of medial caesura declined at the onset of the parallel grand consequent is not applicable. In such cases, the I:HC caesura-gesture is produced too quickly, or the TR-quality of what precedes the I:HC caesura is not sufficiently developed to lead us to expect a real MC so soon in the piece.

When the grand-antecedent paradigm is only slightly expanded, however, the situation becomes more problematic. Suppose that the exposition begins in such a way that we are led to suspect that a recognizable transition (or generic energy-gaining zone) has in fact been entered before the rhetorical I:HC flourish and the subsequent return of P in the tonic. When this occurs, we are confronting a rhetorically stronger, more ambiguous subcategory, which we describe as the *complex grand antecedent*. Consider, for example, the beginning of the *Presto* exposition in the first movement of Haydn's Symphony No. 101 in D ("Clock"). Here we have what seems to be a brief P-idea beginning in m. 24 (itself a terse antecedent-consequent pair) followed at once in m. 34

by the normatively Haydnesque TR-launch through *forte* affirmation, which, in addition, makes a modulatory feint in m. 44 with the introduction of $\sharp 4$, *sforzando*, before the rhetorical flourish at the fermata-sustained I:HC caesura at m. 48.

This situation is more complex than that found in the “Bear” or “Hen” Symphonies. Could this emphatic I:HC punctuation at m. 48 be taken for an MC candidate? Although it does not lead to an S-theme here, launching one was apparently an open possibility: in a few other pieces, Haydn accepts the I:HC as a true MC to produce the early onset of a monothematic S in V. (In the finale of Symphony No. 98 in B \flat , for instance, the I:HC MC is sounded in m. 38 and S, after three hesitant false starts, begins cheerily in F major in m. 43. As it happens, this premature S is immediately cast aside for more TR or *Fortspinnung* rhetoric, as if its early emergence had been a compositional mistake.) More normally, though, Haydn follows a I:HC MC candidate with a restatement of (at least) the opening of P at its original pitch level, thus declining the weak potential MC and eventually proceeding into a new zone of real TR. The resulting implication is that the first apparent TR (before the I:HC caesura) had been only a false transition (the final portion of what we now reconceive as a complex grand antecedent), just as the I:HC MC effect proves to have been a false medial caesura. In mm. 49–50 of the “Clock”—immediately after the I:HC caesura-pause—the grand-consequent residues of the ensuing music are barely perceptible. Only the first nine notes of P are sounded at the original pitch level, and the last two of the nine are reharmonized (m. 50, vii^{o6}/ii). This immediately subverts the P-restatement and moves toward a tonicization of E minor (ii), precipitating the music into more overtly transitional rhetoric, evidently in search of a real MC. (In this instance, it generates a prolonged search that eventually pushes through a V/V structural-dominant lock, beginning in m. 63, to conclude with an emphatic V:PAC MC

in m. 80).³¹ More often, as in the first movement of Symphony No. 94 in G (“Surprise”)—with its auxiliary-cadence P-theme idiosyncrasies—Haydn follows the false I:HC MC with a fuller restatement of P in the tonic, as does Beethoven, for instance, in the first movement of the Piano Sonata in G, op. 31 no. 1. Still, the typical effect in all such cases is that of a complex grand antecedent with false transition leading to a strongly articulated I:HC caesura and a restatement of at least some of P in the tonic (or at the original pitch level, though perhaps reharmonized.) Although the caesura-strength of such a situation varies from case to case, the general principle seems in dialogue with the category of medial caesura declined.

Not all cases of Situation 1 begin by restating P material: sometimes the first MC candidate is followed by a new, lyrical theme (even with typical S-rhetoric) in the tonic key. In such cases the medial-caesura-declined status of the gesture is much clearer. The *locus classicus* occurs in the finale of Beethoven’s Symphony No. 2 in D. (Although this movement is a sonata/rondo intermixture, its expositional principles are those of unmixed sonata form.) Here the PAC that concludes P elides with the onset of a generically normative TR (m. 12) that drives toward an early I:HC caesura at m. 25, one bearing many of the formulaic features of a typical MC: a prolonged dominant, quadruple hammer blows with *Nachschlag*, GP, and so on. Certainly a standard S-theme could emerge at this point, although it would be very early within the exposition to do so. And indeed, the new idea that follows (m. 26, *piano* and lyrical, emerging in cellos and basses) exhibits normative S-behavior in all respects save one: it is solidly anchored in the tonic key. After several broad tonic-dominant oscillations (perhaps suggesting the tonal process of a fugal exposition based upon a modulatory subject), it

³¹This V:PAC and its proportional placement are also discussed in Section 9 below.

moves toward V/V (m. 44), whereupon a new thematic module reinvigorates TR-activity and presses toward a V:HC caesura finally articulated at m. 50. Two measures of caesura-fill lead to a new lyrical theme in the key of the dominant (m. 52, unmistakably S proper).

Situation 2: *Following a proposed MC (usually I:HC or i:HC) the music shifts suddenly onto/into the wrong key.* This produces a tonal *non sequitur*, often suggesting a foreign, flat-side key or chord (♭III, ♭VI, and so on). Moreover, the subsequent music does not proceed efficiently to a PAC in the proper key. (If it does—which would be a rare event in the eighteenth-century style—it might be better interpreted as an S-deformation, one that begins with an off-tonic disturbance, perhaps as the onset of an auxiliary cadence.) Although the passage may begin lyrically (thematically), it usually moves rapidly into transitional or *Fortspinnung* texture, as if to demonstrate its non-S-status and reinforce its impact as a rejection of the proposed MC.

The tonal unexpectedness of this type of declined medial caesura suggests an impulsive “No!” to the preceding caesura. Its precise effect differs according to its circumstances and manner of articulation (lyrical/non-lyrical; loud/soft; and so on). It might suggest a willful, *forte* assertion of personality or eccentricity; a dogged determination not to succumb to a weak caesura; or, conversely, a momentary failure of nerve and tragic slippage onto the wrong key or into a zone of shadowy escape. Generally considered, it suggests either a decisive rejection of the offer to open S-space or a seeming (if temporary) inability to do so.

The first movement of Beethoven’s String Quartet in C minor, op. 18 no. 4, illustrates this general situation (Example 6). At m. 25 the music reaches what at first sounds like an unambiguous i:HC (second-level default) MC, complete with *fortissimo* double hammerstroke and GP. Although the passage that follows (m. 26) might strike us as thematic, there are two reasons that it should not be understood as S proper.

First, launched by an *fp* dynamic shock, the music lurches abruptly here to A♭ major (VI). (True, one might not have expected the sudden appearance of III, E♭ major, directly after a i:HC MC—such is the problem of the second-level-default in minor-mode expositions—but it is hardly inconceivable. Indeed, m. 26 begins with that E♭ pitch sounded in octaves; by m. 27 it is understood, of course, as the unfolded upper fifth of an A♭-major chord.) Second, after four measures, at m. 30, Beethoven begins a sequence of this material a third lower (on F minor or iv), leading ultimately not to a PAC but to a new, less problematic MC candidate, a III:HC (first-level default) at m. 33. This caesura is then accepted as the true MC, and a more normative (unmistakable) S begins in the proper key of E♭ (III) with the upbeat to m. 34, thus finally launching the second part of the exposition. In retrospect the once-potential S of mm. 26–33 can be viewed as a return to the tonal function, if not to the most typical rhetoric, of a transitional zone. The i:HC at m. 25 is probably an applied divider (a backwards-relating dominant). In short, the overall tonal progression in the exposition’s first part is a motion from i (m. 1) through VI (m. 26) to iv (m. 30) to V/III (m. 33). This example also shows a common (though not inevitable) occurrence within medial-caesura-declined situations: as a III:HC, the second MC-candidate is tonally stronger but rhetorically weaker (less rhetorically assertive) than the first.

The first movement of Beethoven’s Violin Sonata No. 5 (“Spring”) in F major, op. 24, offers an even clearer illustration, particularly because additional rhetorical factors complement the unexpected tonal shift. Here the music drives with classic TR-rhetoric toward a I:HC caesura (with energetic, upward-leaping *Nachschlag*) in m. 25. This V of F, however, is followed abruptly by a *fortissimo* scalar unfolding of V⁷–I in A♭ (♭III in the key of F, ♭VI in the key of the coming C major) in mm. 26–28. Even apart from its tonal shift, this gesture is manifestly non-thematic (judged

Example 6. Beethoven, String Quartet in C Minor, op. 18 no. 4, first movement: mm. 20–35

[Allegro ma non tanto]

20 *p*

24 *sf sf ff fp*

28 *fp fp*

32 *p cresc.*

by the norms of standard S-rhetoric): it is non-lyrical; its assertive dynamics are the polar opposite of the normative drop to *piano*; and, if anything, it plunges back into rhetorical activity more generically characteristic of TR-passages. In short, everything about these measures declares an

immediate rejection (“No!”) of the proposed I:HC MC candidate. What follows is a guided path to a second, more acceptable caesura at m. 37: a first-level-default V:HC. Mm. 26–28 initiate a chain of descending fifths ($E\flat-A\flat-D-G-C$), after which an augmented sixth chord on $A\flat$ (m. 33) leads

to the structural dominant in m. 34. The latter is treated to a typical, dynamically reinforced prolongation leading to the MC proper. This includes a hypermetrical reference to the triple-hammer-blow principle (the *Nachschlag*-like g^2-g^1-g in mm. 34–36) followed by a more normative reference to the same principle (*sf* octaves in m. 36 leading to the downbeat of m. 37). Slipping frictionlessly into a standard caesura-fill of the $\hat{5}-\hat{1}$ linear type (linear motion in all voices that sets up the new tonic C with pseudo-cadential effect, as discussed in Section 6 above), this second HC functions as the true medial caesura, and the theme that follows it is S. (We might add that S's agitated accompaniment, *sfp* interjections, and collapse into minor suggest a deformation of generically normative S-behavior. As is common in Beethoven's works, the built-up energy cannot be easily contained, and S threatens to burst through its Enlightenment container.)

Situation 3: *The music following the proposed MC accepts the generically expected new key but decisively reinvigorates obvious TR-texture (or pointedly avoids all features of normative S-rhetoric).* In the above example from Beethoven's "Spring" Sonata, *fortelfortissimo*-underscored TR-activity had been one of the rhetorical signs reinforcing a tonal-shift type of medial caesura declined. Is it possible for emphatic non-S texture alone—occurring within the expected new key—convincingly to decline a proposed MC? This is a difficult question. For the present it seems to us that it sometimes can, although one should be cautious (even skeptical) whenever making such a claim for any individual composition. Each instance must be examined individually, within its own set of tonal and rhetorical circumstances, and in ambiguous cases one can imagine legitimate interpretations differing.

One instance in which a reasonable case for a Situation 3 medial caesura declined may be made occurs in the opening Allegro of Mozart's Symphony No. 20 in D, K. 133 (Example 7). Here TR begins in m. 14 and drives through standard rhetorical signals (for instance, the ongoing *Trommelbass*

figuration outlining the bass motion $\hat{4}-\hat{\#4}-\hat{5}$ in the new key, A major, in mm. 31–32) to manufacture a first-level-default V:HC MC at the first quarter note of m. 34. (One could imagine the remainder of m. 34 as occupied by two quarter rests and an immediate cut to the upbeat of m. 43, which will serve as the true S.) But no caesura-gap appears here. Instead, in m. 34 one experiences a moment of alarm: a refusal to produce a gap or to drop to *piano* dynamics; a continuation of the agitated tremolo in the first violins; and, above all, an instant collapse of the dominant, E-major MC chord (V of A) into minor ("No!"), underscored with an adrenaline-surge swerve (led by the second violins) into further, circle-of-fifths-oriented *Fortspinnung*.³² The expressive effect of all this—surely part of the aesthetic plot devised by the young Mozart—is as if one had been caught off-guard, not fully ready to sound (or to hear) S. Hence the musical impression of suddenly scrambling for recovery. The gesture is all youth and audacity, doubtless a dazzling twisting of the generic tail within the piece's 1772 context.

³²The surprising collapse of the dominant (MC) chord, V/V, into minor is not to be confused with the generically recognizable collapse of the dominant key (V) into minor at the onset of S. Setting up the expectation of a major-mode S but beginning that theme in the dominant minor was not uncommon in the middle decades of the eighteenth century. Charles Rosen discussed this as one of the "three stereotypes of the 1750s and 1760s that were to disappear [in later decades]" and cited some examples in *Sonata Forms*, rev. ed., 153–54. A *locus classicus* from the early 1780s, though not one mentioned by Rosen, occurs in Mozart's Overture to *Idomeneo*, in D major. Here TR seeks ultimately to produce a third-level-default V:PAC MC—one obviously also engaging in a highly sophisticated dialogue with the $\hat{5}-\hat{1}$ caesura-fill principle after an abruptly clipped cadential $\hat{6}_4$ at m. 41—at the downbeat of m. 45. At the moment of PAC-resolution, however, the promised A major is suddenly chilled to A minor. S emerges ominously in that A minor with the upbeat to m. 45; this leads to further rhetorical and tonal adventures before the EEC at the downbeat of m. 64. Another example of a minor-mode S in Mozart occurs in the finale of the Piano Sonata in F Major, K. 332, m. 50.

Example 7. Mozart, Symphony No. 20 in D Major, K. 133, first movement: mm. 30–49

This is also an example of the rarer practice of declining a first-level-default MC, V:HC. One might notice how close conceptually this situation is to the last-possible-moment occurrences of the bait-and-switch tactic. Such resemblances suggest how seamlessly the extreme instances of the bait-and-switch tactic can merge into situations perhaps better considered as medial caesura declined. Here in Symphony No. 20, the new *Fortspinnung* leads to a third-level-default V:PAC in the structural voices (the last-available MC-possibility within the standard deployment-sequence options) and a normative, unmistakable S at m. 43.³³ It is unlikely that

³³At the point of the MC (m. 42), the first oboe plays a c^{#2} (not shown in Example 7). This pitch should be understood as the octave doubling of an

m. 43 should be understood as the onset of C: not only is its rhetoric typical of S, but the exposition as a whole, which ends in m. 78, is only slightly over half completed.

As one confronts non-normative sonata behavior around a proposed MC point, it is easy to overuse the interpretive tool of medial caesura declined. It is a basic proposition of Sonata Theory that if a strong HC-caesura is heard at the S-point and if the following phrase is harmonically and tonally stable in the expected key and mode, that phrase should

inner voice: essentially, the first oboe has been doubling the second violin part an octave higher since the second half of m. 25. Because the first violin's a¹ is the structural top voice of m. 42, this MC should be understood as built around a perfect rather than an imperfect cadence. (Cf. note 15 above.)

normally be understood as S. This seems obvious enough, and problems arise only when what follows the presumed medial caesura is rhetorically unusual—if it begins with sudden *forte* bluster, for instance, instead of the standard drop to *piano*. But such a *forte* surge alone is not always sufficient to justify a claim that the preceding MC candidate has been declined. Sometimes what follows the MC is more judiciously regarded as an *S-deformation*, in which S's normative (first-level-default) rhetoric is overridden for local expressive purposes. When the evidence persuades us that this is the case, the passage should be considered as something in dialogue with the normative S-principle in the proper S-space.

Such an S-deformation (not medial caesura declined) may occur in the initial movement of Mozart's Quartet in C Major ("Dissonance"), K. 465, m. 56.³⁴ Here the transitional zone's articulation of a V:HC medial caesura (m. 55, with GP-gap, though preceded by a curious, non-normative energy-drop to *piano*, in m. 53) is followed by a section of brilliant passagework in the dominant, G major (mm. 56–71). This passage begins with a repeated, short-lived *forte* impulse (2 + 2 mm.—as if trying to compensate, albeit with hesitating, *piano* pullbacks, for the pre-MC energy-drop?), continues in a *Fortspinnung* manner not typical of Mozart's S-themes, and is sealed off at the end with a V:PAC featuring an emphatic trill cadence (m. 71). What follows this V:PAC is a gavotte-like theme in a more Mozart-normative S-rhetoric (mm. 72–79).

However we interpret it, Mozart must have intended us to hear mm. 56–71 as unusual. On the level of rhetorical form (although not on the level of tonal form, since it does proceed in G major) this vigorous passage might tempt one to consider it an example of medial caesura declined—a resumption of quasi-thematic TR-like activity that rejects the weakly pro-

posed, *piano*-articulated MC at m. 55. Such an interpretation is especially attractive because of the more normative S-like theme (mm. 72–79) that directly follows the V:PAC of m. 71. In this interpretation the V:PAC in m. 71 may be considered a third-level-default MC, the V:HC default already having been used up in m. 55; S proper, of course, would begin at m. 72. (The relationship of the V:PAC MC to prior feints at articulating a V:HC MC is discussed in Section 9 below.)

On the other hand, one could also argue—perhaps equally plausibly, given the passage's key, as well as its sentence-format and the presence of an unfilled MC-gap at m. 55—that mm. 56–71 should be regarded as an *S-deformation* in which typically Mozartean, lyrical S-rhetoric has been overridden by *Fortspinnung*-like vigor (blustery, momentarily unsettled, suddenly assertive, and so on, surely in response to the unusually gentle preceding MC). In this interpretation, m. 71 would have to be regarded as the point of the EEC, and one could also suggest that in C1 (m. 72) Mozart retrospectively (valedictorily?) tries to recapture the flavor of the lost (or previously sacrificed) S-rhetoric—as if S-rhetoric had been displaced into C-space.

8. THE MID-EXPOSITIONAL TRIMODULAR BLOCK

As outlined above, situations 1 and 2 of medial caesura declined may begin with a lyrical theme including a drop to *piano*: normative S-rhetoric (though delivered in the wrong key). When such themes exist, they usually dissolve before long into a modulatory passage with TR-rhetoric whose usual function is to generate another, stronger MC candidate. This new caesura is then accepted to open S-space and launch a second, rhetorically normative S (now in the correct key). In these cases the exposition presents us with *apparent double medial caesuras*: two MC-moments are generated because the first one is declined. The first proposed MC (usually a I:HC, as might be expected) comes to be functionally erased by subsequent events; the second is accepted as the real MC,

³⁴Note 35 below discusses a related case, though one following a I:HC MC and seemingly more in dialogue with the principle of the mid-expositional trimodular block (Section 8 below).

the gesture that divides the exposition into two parts and initiates S.

When we confront this thematic subset of medial caesura declined (as in the finale of Beethoven's Second Symphony), our perception of the process staged from the production of the first (rejected) MC candidate through the attainment of the EEC can be that of a single multipartite block set apart from the remainder of the exposition and laid out in *three distinct rhetorical modules*: here, S-like theme (but not S)/ new TR and real MC/real S. The general three-stage pattern to which this situation belongs crops up fairly often within this style. We call it the *mid-expositional trimodular block*.

As a subcategory of medial caesura declined, the type of trimodular block (TMB) presented above is a secondary product of certain kinds of rejection of the initially proposed MC. But this is hardly the most frequent type of TMB encountered in the late-eighteenth-century style. More common is a related situation (not falling under the general category of medial caesura declined) in which both MC candidates lead to S-rhetoric themes in the proper key. This TMB condition occurs when the following elements, or minor variants thereof, are found in succession: 1) the presence of a moment of structural punctuation (usually a GP HC-gap) that seems to be taken (or mistaken) for a proposed MC; 2) the onset of a lyrical theme (S1), usually *piano*, in the proper second key (thus suggesting that the prior HC, however weak or non-normative, has been accepted as an MC);³⁵ 3) the in-

³⁵Alternatively, this proper-second-key material could appear *forte* and blustery. When it does, it generates ambiguities: it could be understood either as an S-deformation or as a situation of medial caesura declined, notwithstanding the appearance of the second key (as outlined at the end of Section 7 above). A *locus classicus* occurs in the first movement of Mozart's Symphony "No. 5" in F, K. 43, where one finds a paradigmatic illustration of apparent double medial caesuras, each of which is sounded as a triple hammer blow followed by a GP-gap (I:HC in m. 13 corrected to V:HC in m. 22). The two apparent MCs are connected with a continued-*forte* passage based on P (m.

ability or unwillingness of S1—sometimes presented as weakened or flawed in some respect—to lead directly to a PAC in the second key (or at least the inability to produce a PAC that is convincing as the EEC); 4) the decay of S1 or its reinvigoration of TR-rhetoric; 5) the setting up of a second MC candidate (usually an HC, since a third-level-default PAC could be interpreted as the EEC); 6) the statement of another theme with S-rhetoric (S2), which now provides the PAC that attains the EEC.

The effect is one of seemingly double S-themes, punctuated in the middle with a second transition and second MC. This circumstance merits attention because it is anything but uncommon in late-eighteenth-century and early-nineteenth-century practice and because it has apparently not been widely recognized as such. Since the first MC opens S-space proper (with its new key), we must regard it, not the second caesura, as the structural MC dividing the exposition into two parts. We understand the second caesura—leading to the second apparent S or S2—as a *postmedial caesura* (PMC). (In other words, a postmedial caesura is an emphatic, MC-like caesura planted within the already-launched second part of a two-part exposition.) One of the functions of the entire strategy, including medial and postmedial caesura and first and second S-themes, is to broaden S-space to cover a larger field of proportional time.³⁶ It is a strategy of expansion, although it can also produce eloquent local effects.

14) and beginning in the proper, though here very insecurely tonicized, key of C major (V). This can strike one as an expressive refusal to open up S-space so soon (again, despite the key), and it quickly redissolves into typical pre-MC rhetoric that shores up and solidifies the modulation to C major. Following the V:HC MC, a much more normatively acceptable S, *piano*, begins in m. 23. However one understands the passage, it is a crystal-clear example of a trimodular block.

³⁶See, however, the caveats in note 44, par. 2, below regarding the frictionless equation of TM modules and labels such as "S1" and "S2."

To summarize: this more common type of TMB encompasses the material produced between the first MC-gesture (the real MC) and the EEC (following S2). The first module (TM1) is lyrical and generally given to S-rhetoric, but it can also give the impression of arriving prematurely, even rashly, as if it had accepted a problematic HC and opened S-space too early (thus predicting a smaller exposition, whose diminutive scale might be out of proportion to its grander rhetoric). Whether or not this is the case, TM1 soon proves unable to produce the EEC. TM1 may itself redissolve into transitional activity or give way to a texture-change at TM2, the module concerned with preparing and/or articulating an emphatic HC in the second key—a postmedial caesura, often more convincing or confident than the earlier MC proper.³⁷ This is followed by the onset of the third module (TM3), a new S (S2, usually begun in classic, *piano* S-rhetoric) that drives toward the PAC and finally accomplishes the EEC.³⁸

³⁷Much more rarely (and more problematically) the postmedial caesura at the end of TM2 may be a third-level-default V:PAC or III:PAC. This occurs in the first movement of Mozart's Piano Quartet in E \flat Major, K. 493, m. 59 (with TM1 beginning at m. 28). The remarks on the PAC caesura in Section 9 below are clearly applicable here. A similar situation exists in the opening movement of Beethoven's Piano Sonata in D Major, op. 10 no. 3, in which a V:PAC postmedial caesura occurs in m. 53. Here, however, the situation is further complicated by a TM1 that opens in B minor (vi, m. 23). The closing comments in the present Section 8 concern a wrong-key TM1.

³⁸The TMB of the type discussed here (in which the two apparent medial caesuras may be interpreted as the MC proper and the PMC) should be distinguished from another common S-strategy, the *multimodular S* (MMS). An MMS features several distinct thematic modules, the last of which effects the EEC. Since the most common number of modules is three, we may also speak of the *trimodular S* (TMS). The various modules of the TMS may be designated as S1, S2, and S3, to distinguish them from the components of the TMB.

Obviously a sentence structure might qualify as a TMS: the presentation, continuation, and cadential phases would correspond to S1, S2, and S3 respectively. But composers (Beethoven especially) sometimes write a certain type of TMS that in some respects parallels the TMB: S1 proves incapable

of effecting a PAC, S2 either temporizes over the dominant or dissolves into ambiguous harmonies, and S3 drives to the cadence. What distinguishes this type of TMS from the TMB is, of course, the absence of a PMC and the general lack of true TR-activity during S2. In addition, S3 (unlike TM3) is characteristically assertive and often *forte*, as though pulling itself together to drive (sometimes in a single cadential phrase) toward the EEC.

Two examples: Mozart, Piano Sonata in D Major, K. 284, first movement (TM1, sentence structure = mm. 22–34; TM2 = mm. 34–38 (beat 1); the decisive TM3 = mm. 38–50); and Beethoven, Piano Sonata in A Major, op. 2 no. 2, first movement (the minor mode [flawed] TM1 = mm. 58–76; TM2 = c. 77–83 [the GP at m. 83 is not a postmedial caesura but the result of a texturally evaporated diminished-seventh chord]; the decisive, major-mode TM3 = mm. 84–92).

Two examples: Mozart, Piano Sonata in D Major, K. 284, first movement (TM1, sentence structure = mm. 22–34; TM2 = mm. 34–38 (beat 1); the decisive TM3 = mm. 38–50); and Beethoven, Piano Sonata in A Major, op. 2 no. 2, first movement (the minor mode [flawed] TM1 = mm. 58–76; TM2 = c. 77–83 [the GP at m. 83 is not a postmedial caesura but the result of a texturally evaporated diminished-seventh chord]; the decisive, major-mode TM3 = mm. 84–92).

³⁹A handy example is available in Philip G. Downs, ed., *Anthology of Classical Music* (New York: W. W. Norton, 1992), 39–45: the opening Presto movement of Stamitz's Symphony in D Major ("La Melodia Germanica," No. 1), c. 1754. Here TR begins at m. 17 and drives within eight measures to a clear (but weak) I:HC MC at m. 24. S begins directly, in A major, at m. 25, effecting premature PACs in mm. 28 and 32. Instead of proceeding to a more properly placed PAC (the EEC), Stamitz drives to a second MC-gesture—this time a stronger one, V:PAC with triple hammer blows, at m. 37. A new S-like melody then emerges in mm. 38 (preparatory material) and 39 (theme proper), leading to the EEC at m. 50; this is followed by reinforcing, post-cadential C-material, mm. 50–58. Thus, TM1 = mm. 25–32; TM2 = mm. 32–37; TM3 (leading to the EEC) = mm. 38–50.

The phenomenon of seemingly double second themes (in which the first S leads to yet another transition and set-up) is mentioned by Eugene K. Wolf, *The Symphonies of Johann Stamitz: A Study in the Formation of the Classic Style* (Utrecht: Bohn, Scheltema & Holkema, 1981), e.g., 151, 199, and 272. Cf. 327–28 ("false transition"). Wolf describes the post-S1 TR-texture as a "secondary transition." On p. 200 he mentions that "this design [including a new, *forte* transition that leads to a second S theme] also appears with some frequency in Viennese symphonies (e.g., by Wagenseil and Dittersdorf)."

default option within expositions.⁴⁰ In the monumentalized, highly personalized style of the late-eighteenth-century masters, the TMB invited the exploration of unusual expressive effects or strong generic deformations, particularly around the area of the MC (sometimes not normatively articulated or stated too early) and the subsequent TM1 (sometimes presented as premature, flawed, or fragile). Within this repertory the concept of the TMB is highly complex, due both to the many ways in which it may be worked out and to the expressive deformations that can accompany individual cases. For the present, we only provide a few examples.

The first movement of Beethoven's Piano Sonata in C Major, op. 2 no. 3, proposes an unequivocal I:HC medial caesura with GP gap in m. 26. What follows is an enormous TMB (with expanded third module) that stretches from m. 27 to the EEC in m. 77. This dramatically eventful block begins with TM1 at m. 27: a generically lyrical, *piano*, S-rhetoric theme, but one that begins in G minor, not in the expected G major (characteristic of the problems often surrounding TM1).⁴¹ To be sure, minor-mode status alone does not disqualify a theme as an S-candidate; such minor-mode onsets for S had been common in the mid-eighteenth century, although this earlier stereotype had declined sharply in the century's last two decades (as discussed in note 32, above). Within the context of this sonata from 1795, however, the

⁴⁰See the first movement, e.g., of J. C. Bach, Harpsichord/Piano Sonata, op. 5 (1766) no. 4 (in E♭), and the finale of the similar sonata, op. 17 (c. 1779) no. 2 (also in E♭).

⁴¹An early version of this theme—also in G minor within a larger C-major first movement—appears in one of Beethoven's earliest chamber works from Bonn, the Piano Quartet in C Major [1785], WoO 36, m. 37. Curiously, in that work the minor-mode theme is the third module of an S that begins normatively in G major (m. 24) with a theme that resembles that of TM3 in the later piano sonata. In short (setting aside a few other complicating factors), the thematic order of the secondary-theme zone in the early Piano Quartet is the reverse of that found in op. 2 no. 3.

sudden collapse into minor is a sign that the arrival of the exposition's second part is stamped with significant anxieties. While still opening S-space through its rhetoric, this sober fall into minor suggests that the weaker I:HC MC was in some way unsatisfactory within the narrative that Beethoven wished to present in this individual sonata.⁴² In the musical story related here, the prior events have produced only an imperfect S, as if the more proper (major-mode) theme were not yet prepared or ready for launching. This flaw, it seems, will have to be expunged through the TMB-strategy, which both permits the generation of a second, remedial S-like theme (S2) and expands the overall scale of the exposition.⁴³

⁴²Because of its initial tonal presentation of the dominant, this is probably not a case of medial caesura declined, although it clearly alludes to that strategy's psychology. Tovey's mild hesitations regarding this passage are instructive. He considered m. 27 as opening the "second group (or transition and second group) in [the] dominant." but his subsequent discussion seems to suggest that the best solution is to consider m. 27 (with its "remarkably modulating theme") the beginning of S (*A Companion to Beethoven's Piano-forte Sonatas* [London: Royal Schools of Music, 1931], 24–25).

⁴³A comparison may be made with much of the material newly added by the pianist, beginning with the G-minor theme (m. 109) in the solo exposition of Mozart's Piano Concerto [No. 21] in C, K. 467, first movement. In effect, the pianist proposes an expansive TMB—functioning broadly as S within the solo exposition—none of whose thematic materials had been heard in the preceding orchestral ritornello: TM1 in G minor (mm. 109–22), eliding impulsively with the MC moment itself and merging into an MC preparation in its final measures; TM2, the articulation of a deliciously drawn-out V:HC postmedial caesura with caesura-fill (mm. 122–27); TM3, the second, more normative S-idea (mm. 128–43), here overriding that one proposed in the orchestral ritornello and leading on its own to the EEC at m. 143. The soloist's TMB S, that is, entirely crowds out the ritornello's proposed S, which appears again only in the recapitulation, where it is relegated to the status of a separate module within the closing zone (m. 351).

From a Schenkerian perspective, the dominant harmony that sets in at m. 103 (and turns to minor at m. 109) is a leaping passing tone (*springender Durchgang*) that gives consonant support to a passing d^2 that moves to the $c\sharp^2$ of mm. 120–21: the augmented-sixth chord that enters here is thus generated by a chromatic voice-exchange ($\begin{smallmatrix} E-D-C\sharp \\ C-G-E\flat \end{smallmatrix}$). Thus all the harmonic events

Beginning in the key of the dominant minor, the troubled TM1 either cannot or chooses not to sustain its G minor, the mark of its imperfection. It begins to modulate sequentially, rising by fifths from G minor to a restatement on D minor (m. 33) to new material on A minor (m. 39). This new *forte* module at m. 39 reinvigorates a more characteristic TR–texture, a common sign of TM2, and leads to a first-level-default (improved) V:HC postmedial caesura at m. 45. Two measures of caesura-fill introduce TM3, now in a sunny, expansive G major, and obviously heard as a corrective counterbalance of the earlier TM1 at m. 47. In typically Beethovenian fashion, securing the EEC turns out to be a strenuous affair. TM3 (or S2) is prolonged through cadence postponement and vigorous textural and thematic shifts (sub-modules within the more generically normative category of TM3). The idea at m. 61, of course, brings back propulsive material from m. 14, and the EEC is attained only with the V:PAC at m. 77.⁴⁴

through m. 121 (the augmented-sixth chord) are under the control of the tonic scale-step; the dominant scale-step enters for the first time at m. 128, with the onset of TM3. (We are indebted to Allen Cadwallader for this tonal reading.)

The TMB strategy within solo expositions—either adding a double-module pair (TM1–TM2, for example, or TM2–TM3) for the solo exposition or creating an entirely new TMB S passage altogether, as here—is something that Mozart found particularly attractive within concertos. Quite obviously, it is an eloquent way of demonstrating the expansive, personally inventive possibilities of the soloist within a still generically recognizable framework.

⁴⁴For some considering this exposition, the question might remain: was m. 27, with its G-minor opening, really the onset of S? (Notice, for example, the extreme skewing of the proportions of the exposition: 26 mm. in the first part, 64 in the second, although that is not uncharacteristic of TMB situations.) Surely there are ambiguities that involve the limits of the procedure of medial caesura declined. But from the perspective of Sonata Theory, which is concerned with uncovering an individual work's dialogue with generic precedent, it would be difficult to maintain that such an emphatically formulaic I:HC caesura with GP–gap at m. 26 (how else are we to hear it?) could be decisively declined by a lyrical, *piano* theme whose only initial flaw is modal,

Recognizing the TMB pattern helps us to deal with expositions that might appear more problematic. In the first movement of Beethoven's Piano Sonata in F Major, op. 10 no. 2, the C-major passage from m. 19 to m. 55 (the EEC) displays the standard features of a TMB: a gap or pause on an HC (m. 18, to which we shall return), ushering in a TM1 with S–rhetoric and in the proper key (mm. 19–30);⁴⁵ a transitional TM2 (mm. 30–36), preparing a postmedial caesura at m. 36 followed by two measures of caesura-fill (mm. 36–37); an expanded TM3 (mm. 38–55), producing the EEC at the end. So much is clear, but there are surrounding problems: the HC preceding the C-major TM1 is a non-normative iii:HC (V of A minor); and that HC, if it is to function as an MC, arrives alarmingly early in the exposition.

With op. 10 no. 2 we are dealing with a purposeful deformation of the norm, not with standard practice. (Again, this is telling: once one is able to articulate the generic norm and its probable limits, one may discern the degree of deformational force that Beethoven applied to that norm for

not tonal, and which, moreover, demonstrates the correction of the problem through a generically recognizable TMB strategy.

In a larger sense, any attempt to map such a label as S1 or S2 directly onto portions of a TMB is to insist on interpreting a more complex expositional phenomenon (the TMB) by means of the conceptual categories of a simpler one (the two-part exposition with non-problematic S). A deeper consideration of m. 27 might produce the conclusion that while TM1 might not be said literally to be S (since S–situations—or situations for which the S–concept was devised in the first place—are normally simpler), it is certainly to be considered as in dialogue with the S–principle. Only in that limited sense may TM1 be said (casually) to be S.

⁴⁵In “Organic Structure in Sonata Form” (1926), Schenker considered this to begin the “antecedent of the second theme”; the module we designate as TM3 he called the “consequent of the second theme.” A translation by Orin Grossman appears in Maury Yeston, *Readings in Schenkerian Analysis and Other Approaches* (New Haven: Yale University Press, 1977), 45–47. Tovey, in *A Companion*, 50, also considered this to be the onset of the “second group.”

special effects.) Here the mood is one of rough impatience. A brief, matter-of-fact P leads to what appears to be the onset of TR at m. 13. (The generic strategy in play is that of producing a dissolving restatement of P.) But that TR is abruptly cast aside before it can expand properly. The result is the sudden iii:HC, whose dominant-chord character and triple-hammer-blow effect seem prematurely to trigger the onset of TM1 (or S1). This is a classic instance of a TM1 that strikes one as rash, as if it had jumped at the first HC that might serve as an MC. This impression of self-indulgent caprice, linked with the rhetorical muscle and generic disorder of the passage, is surely part of the young Beethoven's intended purpose.

Although the above examples may contain elements of uncertainty, more serious ambiguities arise when what would otherwise be considered TM1 appears in the wrong non-tonic key (for example, in the submediant or subdominant within a major-mode exposition). In these cases—the first movement of Beethoven's Piano Sonata in D Major, op. 10 no. 3, provides a particularly challenging example—a TMB pattern overlaps with the situation-2 type of medial caesura declined (outlined in Section 7 above).⁴⁶ The resulting situation occupies a tense conceptual space in dialogue both with medial caesura declined (in which the second caesura is the MC) and with the more standard TMB situation (in which the first caesura is the MC). Trying to remove the ambiguity by declaring on behalf of one interpretation or the other is doubtless to miss the point. In most cases, the composer was probably hoping that we would savor the generic tension of the circumstances at hand, not resolve it.

⁴⁶Among its complications, for instance, is a V:PAC postmedial caesura in m. 53, as discussed in note 37.

9. V:PAC (III:PAC) MEDIAL CAESURAS

Within two-part expositions most medial caesuras are constructed around the half cadence, V:HC and I:HC in major-mode expositions and III:HC (or v:HC) and i:HC in minor-mode expositions. In each case the HC is to be heard as a dramatized sign of preparation for what is to follow. Structural caesuras built around third-level defaults, V:PAC and III:PAC (or, much less commonly, v:PAC), are also readily available, though less frequent options.

The remaining conceptual alternative, a seeming MC articulated by a I:PAC or i:PAC is probably best considered either an extremely low-level MC default (within a somewhat telescoped or abbreviated exposition) or an ad hoc substitution for the MC principle. In eighteenth-century works a I:PAC or IAC leading directly to an obvious S in the new key may occasionally be found in light, small-scale works and in some slow movements. In virtually all cases the PAC or IAC closes off a brief, straightforward P, and the resultant impression is one of omitting the TR-zone altogether. In other words, because of the effective ellipsis of TR, the I:PAC or IAC at the end of P is asked to do double duty as the rhetorical MC. This occurs, for instance, in the first movement of Mozart's Quartet in A Major, K. 169, m. 11, with S beginning in V at m. 12. In larger, more ambitious eighteenth-century works featuring manifest TR-rhetoric at the expected place, the I:PAC or i:PAC MC-effect is rare. Did it not occur in the first movement of Mozart's String Quintet in G Minor, K. 516, we might be disinclined to imagine it as a reasonable possibility for grand-scale compositions.⁴⁷

⁴⁷In this extraordinary exposition, the negative pull of G minor is apparently so strong that TR finds itself unable to pull free of its control. The result is the bleakest MC in the repertory, the i:PAC at m. 29. The S that follows (the rhetorical signals make it clear that this is S) begins in the same, inescapable G minor (m. 30) and finally manages to pull itself up to the proper

On the other hand, V:PACs do occur often enough to warrant more sustained attention. Obviously, PAC MCs are stronger tonal gestures than are HCs. V:PAC caesuras are heard as signs of closure, not of expectancy. Because they alter the expressive norms of the MC, they present special problems of understanding. The tonal task of the exposition is to modulate to the dominant or mediant and to cadence decisively with a satisfactory PAC in this second key. This decisive cadence, we recall, signals the completion of essential expositional closure (EEC). In a two-part exposition this PAC terminates S; tonally, it is S's *raison d'être*. But if TR itself produces a PAC in the second key—before the onset of S—we have a premature arrival of that which is normally reserved to signal the attainment of the EEC. What then is the tonal point of S? Does not this PAC demote S to post-cadential status—to the same status that we normally assign to C?

From a Schenkerian viewpoint, the EEC generally represents the first satisfactory completion of the linear descent (*Zug*)—usually $\frac{5-4-3-2-1}{i-v-i}$ in the second key of the exposition.⁴⁸ (In a major-mode exposition this *Zug* often, although not always, prolongs $\hat{2}_v$ in the tonic key.) If TR terminates in a PAC, the question arises whether this *Zug* has not already been completed. The expected Schenkerian response would be negative, and not without reason: it may well be that the linear progression completed at the end of TR is either unsatisfactory in some way or represents motion into an inner voice, and that the *Kopfton* (for example, $\hat{2}_v$) is still operative as the pieces moves into the S-zone. Whatever the explanation, any such premature PAC in the new key is no matter of relative indifference.

But this matter is yet more complex. From a rhetorical point of view, another relevant matter is how far into an exposition the new key's first PAC is placed. This reopens the thorny question of the limits of the appropriate range for mid-expositional MC-placement, a topic introduced at the end of Section 4 above. Because this proportional question concerning the first PAC in the new key arises with unnerving frequency in Haydn's works, it requires us to acknowledge the problem at the outset and to proceed with caution. The problem restated: emphatic V:PACs in the 55–75% range, more common in Haydn's expositions than one might suspect, can strike one as structurally ambiguous. The later the PAC within this range, the more trenchant the ambiguity: are such V:PACs (or III:PACs), when followed by a shift to differing thematic material, to be understood as MC gestures (followed by a secondary-theme zone) or as early EECs completing the central expansion-section of a continuous exposition?⁴⁹ Here one must consider both the relative lateness of the PAC and the nature of the preceding and following material. Analytical conclusions about individual cases may legitimately differ.

The emphatic, first V:PAC in m. 80 of the first movement of Symphony No. 101 in D major ("Clock"), for example, occurs only 55% of the way through the exposition and is followed by a *piano* theme in V somewhat related to P. Given Haydn's penchant for monothematic or near-monothematic S themes, and given the V:PAC at only the 55% point, one is likely to conclude that the subsequent theme at m. 81 does in fact open the secondary-theme zone. But when a V:PAC MC-candidate is sounded later within an exposition,

⁴⁹Yet another option, though a less common one, would be to follow the mid-expositional V:PAC with multiple (perhaps varied or expanded) reiterations of the cadential module, thus delaying the functional EEC through cadential repetition and producing the second subtype of continuous exposition described in Section 2 above.

mediant major in mm. 36–37 (although further damage to S is also apparent in subsequent measures).

⁴⁸Recall, however, the caveat in note 8 above.

and especially when it leads to essentially new or rhetorically unusual musical material largely unrelated to P, we are confronted with a situation that is more difficult to assess. Such a case occurs in the first movement of Symphony No. 102 in B \flat , which renounces its earlier MC-options to produce an almost comically emphatic V:PAC in m. 80 (the 66% point). This is followed at once by a brief passage of seeming surprise and consternation (mm. 81–92, with V:PAC in m. 92) before producing more obvious closing material (based on a motive from TR) in mm. 92–110. Here it is difficult to ascertain whether mm. 81–92 should be understood as an S–deformation, as an unusual beginning of C (following the early close of an expansion section), or as an ad hoc, non-normative, reactive gesture not adequately described by either of the common labels, S or C. For the present, the question must be left open, to be addressed later, perhaps, in a separate study of Haydn’s engagingly idiosyncratic customizations of more normative sonata practice. We need only register here the frequency with which Haydn lays down a V:PAC (or III:PAC) in the 60–75% range and follows it with contrasting material that could be interpreted as either S or C (in either case, more than a little uneasily) within the conventional system of categories.

Let us return to those situations in which we may be more comfortable in assigning S–status to the music following a V:PAC (or III:PAC) MC. In order for us to be convinced that such an S has not lost its cadential status (its tonal birthright, so to speak), we must agree to understand that the PAC at the end of TR is provisional, not binding. It does not stick as a convincing EEC, and the onset of S should be understood as undoing or re-opening its patent (though secondary) effect of closure. One way of dealing with this would be to propose that such an MC, V:PAC (or III:PAC), declares its own prematurity by occurring within the TR–zone (albeit as its final gesture). This would mean that in a two-part exposition the EEC, by definition, must occur within the S–zone. Accepting

this proposition—and we do accept it—requires that one be able to identify intended S–zones on the basis both of characteristic rhetoric and of proportional placement within the exposition.

Perhaps another reason that PAC medial caesuras are structurally weaker than later, EEC-gestures is that they often give the impression of being closely related to—perhaps even highly developed instances of—the procedure of (post-HC) caesura-fill of the $\hat{5}$ – $\hat{1}$ -descent type (discussed in Section 5 above). In this caesura-fill type a V:HC (or III:HC) MC is bridged with a linear descent that connects the dominant of the MC to the tonic of S, as if setting up the new-tonic platform from which S emerges. When this descent is thickened to include all voices (as mentioned earlier), it can give the impression of a V:PAC (or III:PAC) at the moment of S–launch. All that is needed to reinforce this impression—to produce an unequivocal PAC—is to emphasize this cadential quality rhetorically (for example, to produce it aggressively, *forte*, or to reinforce it with a trill cadence or another cadential formula) and, in most cases, to separate the PAC from the ensuing S with a GP–gap. In short, many PAC medial caesuras are preceded with a feint toward an earlier, more normative V:HC MC, a feint that then—in dialogue with the expanded $\hat{5}$ – $\hat{1}$ caesura-fill principle—seems to change its mind in order to produce a more decisive PAC instead. (The effect is often: “No! This time let’s go directly for the authentic cadence!” On the other hand, it may be that a $\hat{5}$ – $\hat{1}$ caesura-fill expansion carried out too far or pushed beyond a certain level of expressive tension demands a stronger resolution by means of a PAC.)

The first movement of Mozart’s String Quartet in D Minor, K. 421, is paradigmatic. Here S begins in F major, III, at m. 25, but it is prepared by a III:PAC MC in the preceding measure. Only a few measures earlier, however, TR had arrived at a strong HC (m. 18, beats 3–4) and reinforced it with two repetitions in the following two measures (mm. 19–

20). By m. 20 TR has virtually produced the expected III:HC MC, but that effect is overridden with an unexpected *forte* gesture in the cello (“No! Let’s do something different!”) that initiates a new surge in all voices, leading to the III:PAC MC at m. 24 (about 60% of the way through the exposition). The closeness of mm. 21–24 to the $\hat{5}$ – $\hat{1}$ caesura-fill principle seems obvious, and Mozart underscored it by calling for an energy-loss in those measures (notice the *piano* dynamic in m. 23, not normally characteristic of standard drives toward an MC) and by following it with the chattering accompanimental pattern in the last half of m. 24 (itself caesura-fill) that bridges the MC to S. The S that follows is brief (mm. 25–32, a repeated sentence built around an auxiliary cadence): it displays few, if any, signs of having been disturbed by the preceding PAC.

In some instances, however, the subsequent S does seem markedly unsettled in these circumstances. In the first movement of Mozart’s Symphony No. 36 in C, K. 425 (“Linz”), the end of the TR is signalled by a V:PAC MC in m. 71. This V:PAC—sounded, as it turns out, at the 52% point through the exposition—is preceded by very curious activity: a standard TR–*forte* drive that breaks off suddenly on a predominant chord (locally, a ii⁶) before the production of the expected MC (m. 62, a blocked-caesura situation that presents a clear exception to normative medial-caesura practice); and a *piano*, caesura-fill-like texture (energy-loss) that itself begins to suggest V:HC caesura-like signals (approach via #4, neighboring motion, and so on, touching significantly on the structural dominant at m. 66). Perhaps Mozart suggests here that V:HC MCs ought not—cannot?—be produced within a texture of energy-loss (the premature caesura-fill texture). The result is that he undercuts his potential V:HC medial caesura and leads the music into a V:PAC at m. 71.⁵⁰

⁵⁰Similar situations are mentioned in note 18 within Section 4 above.

However we interpret these TR–disturbances, they do provoke a reaction from S (an S–deformation), which bursts in with a non-normative show of bluff force: a *forte* Turkish march (m. 72) beginning off-tonic in the key of E minor (vi/V). This antecedent phrase is tamed by a *piano*, buffa-like consequent that begins (m. 76) on C major (IV/V but also an oblique reference to the original tonic key) and cadences in G (V:PAC). The harmonic strategy of S is clear: the overall vi–IV–V–I progression represents an auxiliary cadence in the key of the dominant, a forceful way of re-opening and then reclosing the previous V:PAC. The EEC issue is re-opened once again by a standard repetition of the S–theme (mm. 80–87), reversing the dynamic levels of the two phrases. The EEC (and the end of S) occurs on the downbeat of m. 87—the only time that the V:PAC is delivered *forte*—at which point the music enters a closing zone.

IO. EPILOGUE: BRIEF HISTORICAL CONSIDERATIONS

Sonata form was an ordered system of generically available options permitting the spanning of ever larger expanses of time. A sonata-form composition was a feat of engineering, like the construction of a bridge thrown out into space. In the eighteenth-century style, this temporal span was to be built from rather simple materials: trim, elementary musical modules whose very brevity and small-scale balances seemed best suited to short-winded compositions. One of the central challenges facing the mid- and late-eighteenth century was to use this seemingly unassuming, *galant* language, grounded in structural punctuation and periodicity, to produce ever more spectacular spans for occasions of special dignity, prestige, or social importance. Multimovement conventions, ever-larger, thematically differentiated binary structures (with built-in repetitions), and slow introductions all had their roles to play in this process of generic enlargement. And ultimately they

led to the grandly monumental, personalized structures of Haydn, Mozart, and Beethoven.

Within what was eventually called the exposition of a sonata form, the mid-expositional relaunch into a separately articulated, contrasting second part (S-theme), came to be a crucial strategy of this spanning-procedure. (The main alternative, less often used with every decade in the eighteenth century, was that of an ongoing *Fortspinnung*.) This second-launch strategy was anticipated in the initial sections of the da capo arias of Leonardo Vinci in the 1720s, and, within instrumental compositions, it emerged clearly in the late 1730s and 1740s with the dramatized first movements of certain Neapolitan opera overtures by such successors to Vinci as Leo and Jommelli.⁵¹ It soon became a standard feature of certain types of expositions within early sonata, chamber, and symphonic practice.⁵²

The compositional device that made this second launch possible was the medial caesura, along with the transitional energy-gain leading up to it. Thus it is no surprise that a conventionalized set of options concerning its treatment emerged and continued to multiply at mid-century. It is this set of gradually transforming norms or defaults that young

⁵¹Kurt Sven Markstrom, "Burney's Assessment of Leonardo Vinci," *Acta musicologica* 67 (1995): 142–63 (especially 153–54, which contain references to the anticipation of sonata form).

⁵²Eugene K. Wolf discusses the origins of separate S-themes (thematic differentiation) within more or less expansive binary structures in "Sonata Form," *The New Harvard Dictionary of Music*, ed. Don Michael Randel (Cambridge: Harvard University Press, 1986), especially 765–66; and, in more detail, Helmut Hell, *Die neapolitanische Opernsinfonie in der ersten Hälfte des 18. Jahrhunderts* (Tutzing: Schneider, 1971).

Haydn and, later, young Mozart (as well as every other composer of the period) assimilated and customized into their own personal styles.

By the later eighteenth century, especially in the hands of Haydn, Mozart, and Beethoven, the variety of ways in which the crucial MC could be articulated, 'suggested, evaded, masked, declined, bridged, or stretched had grown enormously. What had arisen as a simple, straightforward device of *galant* structural punctuation—a gap of silence—came to be among the most central, most subtle, most expressive features of the late-eighteenth- and early-nineteenth-century exposition. Sonata Theory, concerned with recreating the tacit dialogue between an individual work and a relevant, historically fluid set of generic norms, seeks to provide the tools for awakening the expressive impact of such things as the inexhaustibly varied treatments of the all-important medial caesura.

ABSTRACT

The authors are developing a genre-based "Sonata Theory," according to which moment-to-moment compositional choices in sonata-form works are understood as elements of an ongoing dialogue with reasonably ascertainable, flexible norms. One of the most important tonal/rhetorical features of the "two-part exposition" (i.e., an exposition with a secondary theme) c. 1800 is the *medial caesura*, bringing an emphatic end to Part One and simultaneously making available the "secondary-theme zone" that launches Part Two. Composers treated the medial caesura in several standard ways, but they could also subject it to generic "deformation" for structural or expressive reasons that become clearer once the norms surrounding medial-caesura activity are understood.