Martin Rohrmeier

The Syntax of Jazz Harmony: Diatonic Tonality, Phrase Structure, and Form

Abstract

The regularities underlying the structure building of chord sequences, harmonic phrases, and combinations of phrases constitute a central research problem in music theory. This article proposes a formalization of Jazz harmony with a generative framework based on formal grammars, in which syntactic structure tightly corresponds with the functional interpretation of the sequence. It assumes that chords establish nested hierarchical dependencies that are characterized by two core types: preparation and prolongation. The approach expresses diatonic harmony, embedded modulation, borrowing, and substitution within a single grammatical framework. It is argued in the second part that the proposed framework models not only core phrase structure, but also relations between phrases and the syntactic structures underlying the main forms of Jazz standards. As a special case, the Blues form relies heavily on the plagal derivation from the tonic and is analyzed in comparison with other analytical approaches to the Blues. The proposed theory is specified to a sufficient level of detail that it lends itself to computational implementation and empirical exploration, and this way it makes a step towards music theory building that embraces the close links between formal, mathematical, and computational methods.

Keywords

music; jazz; syntax theory; harmony; music theory; generative modeling

MUSIC THEORY & ANALYSIS

International Journal of the Dutch-Flemish Society for Music Theory

VOLUME 7, # 1, APRIL 2020, 1–62

https://doi.org/10.11116/MTA.7.1.1

The Syntax of Jazz Harmony: Diatonic Tonality, Phrase Structure, and Form

Martin Rohrmeier

"All musicians are subconsciously mathematicians."
– Thelonious Monk, interview *Down Beat*, 28 October 1971[1]

1. Introduction

A central question in music theory concerns the understanding of the rules behind harmonic sequences, or, in other words, of the underlying logic regulating the dependencies between chords and the overarching coherence of harmonic sequences. There are many theories of harmony[2] and also of Jazz harmony specifically.[3] However, despite a substantial body of literature, challenges and open questions remain in the theory of (Jazz) harmony. First, while many theoretical accounts are mainly focused on the identification of families of chords or single chord relations, the characterization of the principles of the structure building of larger musical phrases and of dependencies governing chord sequences with formally concise models is still barely explored. A second challenge comes from extended tonality and its harmonic implications. Music of the nineteenth century up to Jazz explores a whole range of new harmonic options, relations, and principles that still require further music-theoretical work.

In music, structure and experience are tightly linked. Therefore, music theory lies at a philosophically and methodologically challenging intersection between psychology and

1 John Bartlett, *Bartlett's Familiar Quotations*, Eighteenth Edition (2012).

2 Stefan Kostka and Dorothy Payne, *Tonal Harmony with an Introduction to 20th-Century Music* (New York: McGraw-Hill, 1984); Edward Aldwell and Carl Schachter, *Harmony and Voice Leading*, 2nd ed. (San Diego: Harcourt Brace Jovanovich, 1989); Robert Gauldin, *Harmonic Practice in Tonal Music* (New York: Norton, 1997).

3 Mark Levine, *The Jazz Theory Book* (Petaluma, CA: Sher Music, 1996); Mark Levine, *The Jazz Piano Book* (Petaluma, CA: Sher Music, 1990); Wolf Burbat, *Die Harmonik des Jazz* (Kassel: Bärenreiter, 1988); Frank Sikora, *Neue Jazz-Harmonielehre: Verstehen, Hören, Spielen; von der Theorie zur Improvisation* (Mainz: Schott Music, 2017); Joe Mulholland and Tom Hojnacki, *The Berklee Book of Jazz Harmony* (Hal Leonard Corporation, 2013); Bill Dobbins, *A Creative Approach to Jazz Piano Harmony* (Rottenberg, Germany: Advance Music, 1994); Dariusz Terefenko, *Jazz Theory: From Basic to Advanced Study* (New York: Routledge, 2014).

Music Theory & Analysis | Volume 7, # I, April 2020 | 1–62

| https://doi.org/10.1116/MTA.7.1.1

formal modeling because it links human musical experience with the formal/mathematical characterization of the structures underlying it. Hence, music theory lies at the heart of the research program of the Cognitive Sciences.[4] Music-analytical and theoretical work reveals implicit assumptions, makes implicit expert knowledge[5] explicit, and explains such knowledge through formal tools. In the process of theory building, the listening experience of certain features is sharpened, modeled, and explicated through formal objects and their relations.[6] For this purpose, the precision of mathematical formalization avoids vagueness, ambiguity, and misunderstandings in the theories proposed. Formal theory building is thus closely linked with mathematical and computational modeling[7], since the essence of computation lies in abstraction, functional modeling and function composition.[8]

When it comes to building formal theories of music, specifically musical *sequence*, the theoretical branch of *generative modeling* is particularly relevant. The core of generative

4 Marcus Pearce and Martin Rohrmeier, "Music Cognition and the Cognitive Sciences," *Topics in Cognitive Science* 4, no. 4 (2012): 468–484; Ray Jackendoff and Fred Lerdahl, "The Capacity for Music: What is It, and What's Special About It?," *Cognition* 100, no. 1 (2006): 33–72; Uwe Seifert, *Systematische Musiktheorie und Kognitionswissenschaft: Zur Grundlegung der kognitiven Musikwissenchaft*, vol. 69 (Bonn: Verlag Für Systematische Musikwissenschaft, 1993); Stefan Koelsch, *Brain and Music* (Oxford: John Wiley & Sons, 2012).

5 John A. Sloboda, "Musical Expertise," in *Toward a General Theory of Expertise: Prospects and Limits*, ed. Anders Ericsson and Jacqui Smith (Cambridge: Cambridge University Press, 1991), 153–171; Craig Speelman, "Implicit Expertise: Do We Expect Too Much From Our Experts," in *Implicit and Explicit Mental Processes*, ed. Kim Kirsner et al. (New Jersey: Erlbaum, 1998), 135–147; Arthur S. Reber, *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious* (Oxford: Oxford University Press, 1993); Barbara Tillmann, "Implicit Investigations of Tonal Knowledge in Nonmusician Listeners," *Annals of the New York Academy of Sciences* 1060, no. 1 (2005): 100–110; Martin Rohrmeier and Patrick Rebuschat, "Implicit Learning and Acquisition of Music," *Topics in Cognitive Science* 4, no. 4 (2012): 525–553; David Huron, *Sweet Anticipation: Music and the Psychology of Expectation* (Cambridge, MA: MIT Press, 2006).

6 Ian Cross, "Music Analysis and Music Perception," *Music Analysis* 17, no. 1 (1998): 3–20; Huron, *Sweet Anticipation: Music and the Psychology of Expectation*; Markus Neuwirth and Martin Rohrmeier, "Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung," *Zeitschrift der Gesellschaft für Musiktheorie* 13, no. 2 (2016): 171–193, doi:https://doi.org/10.31751/915; Martin Rohrmeier, "Musical Expectancy: Bridging Music Theory, Cognitive and Computational Approaches," *Zeitschrift der Gesellschaft für Musiktheorie* 10, no. 2 (2013): 343–371; Geraint A. Wiggins, "Music, Mind and Mathematics: Theory, Reality and Formality," *Journal of Mathematics and Music* 6, no. 2 (2012): 111–123, doi:10.1080/17459737.2012.694710; David Huron, *Voice Leading: The Science Behind a Musical Art* (Cambridge, MA: The MIT Press, 2016); Fred Lerdahl, *Composition and Cognition: Reflections on Contemporary Music and the Musical Mind* (Oakland, California: University of California Press, 2020).

7 Geraint Wiggins, "Computational Models of Music," in *Music and Language as Cognitive Systems*, ed. Patrick Rebuschat et al. (Oxford: Oxford University Press, 2011), 169–188; David Temperley, "Computational Models of Music Cognition," in *The Psychology of Music*, ed. Diana Deutsch (Amsterdam: Elsevier, 2012), 327–368; David Temperley, "The Question of Purpose in Music Theory: Description, Suggestion, and Explanation," *Current Musicology* 66 (1999): 66–85; Martin Rohrmeier and Marcus Pearce, "Musical Syntax I: Theoretical Perspectives," in *Springer Handbook of Systematic Musicology*, ed. Rolf Bader (Berlin-Heidelberg: Springer, 2018), 473–486.

8 Bartosz Milewski, *Category Theory for Programmers* (London, UK: Blurb Books, 2018); Michael Barr and Charles Wells, *Category Theory for Computing Science* (New York: Prentice Hall, 1990); Neil Jones, *Computability Theory: An Introduction* (New York, NY: Academic Press, 1973); Alonzo Church, "An Unsolvable Problem of Elementary Number Theory," *American Journal of Mathematics*, no. 58 (1936): 34–363.

thinking lies in the analysis of a given phenomenon in terms of the principles underlying its (re-)construction (i.e. *generation*) under assumed object relations. For instance, formal grammars allow complex structures to be modeled as being generated from the (recursive) recombination of simple primitives, such as constitutive elements and rules, which together characterize the dependency structures underlying the phenomenon (as employed in this article). The frameworks of "generative syntax" or "generative modeling" are closely tied to the Chomskian tradition of syntactic research[9], but are also deeply connected to the branches of Bayesian cognitive modeling, generative statistical models and Bayesian theory of science.[10] Crucially, generative competence (as the capacity to re-construct a sequence with a generative model) is not the same as the ability to fully compose music or speak a language, since there is a fundamental difference between the ability to generate or analyze the categories, dependencies, and constituents within a well-formed sentence and having something to express (which involves much more, such as semantic coherence, intentions, internal mental states, and pragmatic world-knowledge). A formal generative theory of competence models the former but not the latter. An understanding of the generative principles of music and a model of the musical composition process are conceptually distinct.

Since its origins, generative modeling has developed in numerous directions. In particular, the combination of generative models with probabilistic modeling has proved highly successful in the Cognitive Sciences when modeling the inference of complex structured representations from a world replete with uncertainty and messy conditions. Generative modeling may also involve various kinds of structured models; it does not necessarily relate to grammars or hierarchical, tree-based models. Overall, such directions involve (probabilistic) models of causal inference, rational reasoning, language, vision, object recognition, decision making, action planning, intuitive physics, and music, among others.[11]

---

9 E.g. Noam Chomsky, *Syntactic Structures* (The Hague: Mouton, 1957); Noam Chomsky, *Aspects of the Theory of Syntax* (Cambridge, MA: The MIT Press, 1965); Noam Chomsky, *The Minimalist Program* (Cambridge, MA: The MIT Press, 1995); David Adger, *Core Syntax: A Minimalist Approach* (Oxford: Oxford University Press, 2003).

10 Thomas L. Griffiths, Charles Kemp, and Josh B. Tenenbaum, "Bayesian Models of Cognition," in *The Cambridge Handbook of Computational Cognitive Modeling*, ed. Ron Sun (Cambridge: Cambridge University Press, 2008); Edwin T. Jaynes, *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press, 2003); David J. C. MacKay, *Information Theory, Inference, and Learning algorithms* (Cambridge: Cambridge University Press, 2003).

11 E.g. Griffiths, Kemp, and Tenenbaum, "Bayesian Models of Cognition"; Nick Chater, Josh B. Tenenbaum, and Alan Yuille, "Probabilistic Models of Cognition: Conceptual Foundations," *Trends in Cognitive Sciences* 10, no. 7 (2006): 287–291; Joshua B. Tenenbaum, Thomas L. Griffiths, and Charles Kemp, "Theory-based Bayesian Models of Inductive Learning and Reasoning," *Trends in Cognitive Sciences* 10, no. 7 (2006): 309–318, doi:10.1016/j.tics.2006.05.009; Joshua B. Tenenbaum et al., "How to Grow a Mind: Statistics, Structure, and Abstraction," *Science* 331, no. 6022 (2011):

When it comes to the specific problem of modeling sequential structure, such as harmonic sequences in music, the model class of *formal grammars* or *generative grammars* lends itself particularly well to the problem.[12] Formal grammars constitute a powerful resource for music theory, and their application already has a tradition.[13]

One line of research concerns strictly local grammars and their computational probabilistic counterparts, Markov and n-gram models, being applied to modeling melodic structure, voice-leading, and harmony:[14] Strictly local models characterize sequential structure by modeling the transitions between single elements. *Computation* as a theoretical principle precedes computers[15], and in this sense Piston's table of common root progressions[16] constitutes an oft-cited intuitive version of a generative Markov model, as might be similarly argued for the transitions specified by Rameau's theory of the *basse fondamentale*[17]. Often, Markov tables are used to characterize one-to-one

---

1279–1285, ISSN: 0036-8075, doi:10.1126/science.1192788; Mike Oaksford and Nick Chater, *Bayesian Rationality: The Probabilistic Approach to Human Reasoning* (Oxford: Oxford University Press, 2007); Brian Christian and Tom Griffiths, *Algorithms to Live By: The Computer science of human decisions* (London: Macmillan, 2016).

12 Michael Sipser, *Introduction to the Theory of Computation* (Boston, MA: Cengage Learning, 2012); Uwe Schöning, *Theoretische Informatik—Kurz Gefasst* (Mannheim: Wissenschaftsverlag, 1992); Gerhard Jäger and James Rogers, "Formal Language Theory: Refining the Chomsky Hierarchy," *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, no. 1598 (2012): 1956–1970.

13 See Rohrmeier and Pearce, "Musical Syntax I: Theoretical Perspectives", for a review.

14 Moray Allan and Christopher Williams, "Harmonising Chorales by Probabilistic Inference," in *Advances in Neural Information Processing Systems (NIPS)*, ed. Y. Weiss, B. Schölkopf, and J.C. Platt, 18 (2005), 25–32; Marcus T. Pearce, "The Construction and Evaluation of Statistical Models of Melodic Structure in Music Perception and Composition" (PhD diss., City University London, 2005); Alastair Craft and Ian Cross, "A N-gram Approach to Fugal Exposition Composition," in *Proceedings of the AISB Symposium on Artificial Intelligence and Creativity in the Arts and Sciences* (2003), 36–41; Marcus Pearce and Martin Rohrmeier, "Musical Syntax II: Empirical Perspectives," in *Springer Handbook of Systematic Musicology*, ed. Rolf Bader (Heidelberg: Springer, 2018), 487–505; Rohrmeier, "Musical Expectancy: Bridging Music Theory, Cognitive and Computational Approaches"; Darrell Conklin and Ian H. Witten, "Multiple Viewpoint Systems for Music Prediction," *Journal of New Music Research* 24, no. 1 (1995): 51–73; Christopher Antila and Julie Cumming, "The VIS Framework: Analyzing Counterpoint in Large Datasets," in *Proceedings of the 15th International Society for Music Information Retrieval Conference, ISMIR 2014*, ed. Hsin-Min Wang, Yi-Hsuan Yang, and Jin Ha Lee (Taipei, Taiwan, 2014), 71–76, Raymond P. Whorley and Darrell Conklin, "Music Generation from Statistical Models of Harmony," *Journal of New Music Research* 45, no. 2 (2016): 160–183; Dan Ponsford, Geraint Wiggins, and Chris Mellish, "Statistical Learning of Harmonic Movement," *Journal of New Music Research* 28, no. 2 (1999): 150–177; Peter Schubert and Julie Cumming, "Another Lesson from Lassus: Using Computers to Analyse Counterpoint," *Early Music* 43, no. 4 (2015): 577–586; Darrell Conklin, "Modelling and Sampling Jazz Chord Sequences," in *Proceedings of the 10th International Workshop on Machine Learning and Music, 6.10. 2017, Barcelona, Spain* (2017), 13–18; Christoph M Wilk and Shigeki Sagayama, "A Parameterized Harmony Model for Automatic Music Completion," *Journal of Information Processing* 28 (2020): 258–266, doi:10.2197/ipsjjip.28.258.

15 William Aspray, ed., *Computing before Computers* (Ames: Iowa State University Press, 1990).

16 Walter Piston, *Harmony* (New York: W.W. Norton & Company, 1948).

17 Jean-Philippe Rameau, *Traité de l'harmonie reduite à ses principes naturels: divisé en quatre livres* (Ballard, 1722); Thomas Hedges and Martin Rohrmeier, "Exploring Rameau and Beyond: A Corpus Study of Root Progression Theories," in *International Conference on Mathematics and Computation in Music* (Springer, 2011), 334–337; Rohrmeier, "Musical Expectancy: Bridging Music Theory, Cognitive and Computational Approaches."

transitions between chords, such as in Bach's Chorales[18], Beethoven's string quartets[19], and Jazz or Rock pieces[20].

Context-free grammars and derived formalisms go beyond strictly local models insofar as they are less restricted in the kinds of dependencies they can express: they are capable of expressing not only local, but also nonlocal dependencies between elements, as well as the hierarchical grouping of elements that belong together (constituency structure); as such they are better suited tools for music theorists' expressive requirements than local grammars. Analyses with such kinds of models may be expressed with tree structures (parse trees), which depict the generative derivation of the sequence. Leonard Bernstein's famous Norton lectures at Harvard in 1973 took inspiration from the recent generative linguistics[21], and among the first to propose to model harmonic and melodic structure with context-free grammars and related ideas were Allan Keiler, Mario Baroni, Mark Steedman, and others.[22] Taking intuition from linguistics to music theory, Fred Lerdahl and Ray Jackendoff published a highly influential theory of metrical, grouping, and reductive analysis of tonal music (henceforth GTTM[23]), drawing from the principles of

18 Martin Rohrmeier and Ian Cross, "Statistical Properties of Tonal Harmony in Bach's Chorales," in *Proceedings of the 10th International Conference on Music Perception and Cognition* (2008), 619–627; Dmitri Tymoczko and Nicolas Meeùs, "Progressions Fondamentales, Fonctions, Degrés: Une Grammaire de l'Harmonie Tonale Élémentaire," *Musurgia* 10, nos. 3–4 (2003): 35–64.

19 Fabian C. Moss et al., "Statistical Characteristics of Tonal Harmony: A Corpus Study of Beethoven's String Quartets," *PloS one* 14, no. 6 (2019).

20 Martin Rohrmeier and Thore Graepel, "Comparing Feature-based Models of Harmony," in *Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval* (2012), 357–370; Trevor De Clercq and David Temperley, "A Corpus Analysis of Rock Harmony," *Popular Music* 30, no. 1 (2011): 47–70; Yuri Broze and Daniel Shanahan, "Diachronic Changes in Jazz Harmony: A Cognitive Perspective," *Music Perception: An Interdisciplinary Journal* 31, no. 1 (2013): 32–45; Klaus Frieler et al., "Two Web Applications for Exploring Melodic Patterns in Jazz Solos," in *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR)* (2018), 777–783.

21 Leonard Bernstein, *The Unanswered Question: Six Talks at Harvard* (Cambridge, MA: Harvard University Press, 1976).

22 Allan Keiler, "Bernstein's 'The Unanswered Question' and the Problem of Musical Competence," *The Musical Quarterly* 64, no. 2 (1978): 195–222; Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid* (London: Penguin, 1979); Curtis Roads, "Grammars as Representations for Music," *Computer Music Journal*, no. 3 (1 1979): 48–55; Mario Baroni, Simon Maguire, and William Drabkin, "The Concept of Musical Grammar," *Music Analysis* 2, no. 2 (1983): 175–208; M. Baroni et al., "A Grammar for Melody: Relationships Between Melody and Harmony," in *Musical Grammars and Computer Analysis*, ed. Mario Baroni and Laura Callegari (Firenze: Leo S. Olschki Editore, 1982); Mario Baroni and Carlo Jacoboni, "Analysis and Generation of Bach's Chorale Melodies," in *Proceedings of the First International Congress on the Semiotics of Music, Centro di Iniziativa Culturale, Pesaro, Italy* (1975); Mario Baroni and Carlo Jacoboni, "Computer Generation of Melodies: Further Proposals," *Computers and the Humanities*, 1983, 1–18; Mark Steedman, "A Generative Grammar for Jazz Chord Sequences," *Music Perception* 2, no. 1 (1984): 52–77; Mark Steedman, "The Blues and the Abstract Truth: Music and Mental Models," chap. 15 in *Mental Models in Cognitive Science: Essays in Honour of Phil Johnson-Laird*, ed. Jane Oakhill and Alan Garnham (Hove, UK: Psychology Press, 1996), 305–318; Tim Horton, "The Formal Structure of Tonal Theory" (PhD diss., Faculty of Music, University of Cambridge, 2002).

23 Fred Lerdahl and Ray Jackendoff, *A Generative Theory of Tonal Music* (Cambridge, MA: The MIT Press, 1983).

Schenkerian theory[24]. However, their framework does not constitute a formal generative grammar, because they do not specify generative rules (or generative principles of concise top-down forward generation) and propose instead a bottom-up analysis that builds up a tree representation (the formulability of which in terms of a corresponding top-down formal grammar, as well as its context-freeness, are unclear and not necessarily guaranteed). Schenkerian theory itself has also found several applications to Jazz, particularly in the work of Steve Larson, Henry Martin, and others.[25] Working at a note-level representation, there is a recent line of hierarchical Schenkerian-like reductions employing *maximal outerplanar graphs*.[26]

Generative approaches to musical grammars have seen a number of remarkable developments. Mark Steedman developed a formal grammar of the 12-bar blues.[27] Satoshi Tojo proposed a Head-Driven Phrase Structural Grammar (HPSG) model[28] of tonal cadences[29]. I myself developed a generative syntax model of tonal harmony (GSM), which employs symbols amended with features like key.[30] Markus Neuwirth and I extended the GSM to characterize the complexity of the variety of tonal cadences, including perfect,

24 Heinrich Schenker, *Der Freie Satz (= Neue musikalische Theorien und Phantasien 3)*, 2nd ed., ed. Oswald Jonas (Vienna: Universal Edition, 1956); Allen C. Cadwallader and David Gagné, *Analysis of Tonal Music: A Schenkerian Approach* (New York: Oxford University Press, 2007).

25 Steve Larson, *Analyzing Jazz: A Schenkerian Approach* (Hillsdale, NY: Pendragon Press, 2009); Steve Larson, "Musical Forces, Melodic Expectation, and Jazz Melody," *Music Perception: An Interdisciplinary Journal* 19, no. 3 (2002): 351–385; Steve Larson, "Schenkerian Analysis of Modern Jazz: Questions About Method," *Music Theory Spectrum* 20, no. 2 (1998): 209–241; Henry Martin, "Schenker and the Tonal Jazz Repertory," *Dutch Journal of Music Theory* 16, no. 1, 1–20; David J. Heyer, "Applying Schenkerian Theory to Mainstream Jazz: A Justification for an Orthodox Approach.," *Music Theory Online* 18, no. 3 (2012); Benjamin Givan, "Swing Improvisation: A Schenkerian Perspective," *Theory and Practice* 35 (2010): 25–56; Mark McFarland, "Schenker and the Tonal Jazz Repertory: A Response to Martin," *Music Theory Online* 18, no. 3 (2012); Dariusz Terefenko, "Keith Jarrett's Art of Solo Introduction: 'Stella by Starlight'—A Case Study," *Intégral* 24 (2010): 81–114; Richard S. Pellegrin, "On Jazz Analysis: Schenker, Salzer, and Salience" (PhD diss., University of Washington, 2013).

26 Jason Yust, *Organized Time: Rhythm, Tonality, and Form* (Oxford: Oxford University Press, 2018); Phillip B. Kirlin and Jason Yust, "Analysis of Analysis: Using Machine Learning to Evaluate the Importance of Music Parameters for Schenkerian Analysis," *Journal of Mathematics and Music* 10, no. 2 (2016): 127–148; Christoph Finkensiep, Richard Widdess, and Martin Rohrmeier, "Modelling the Syntax of North Indian Melodies with a Generalized Graph Grammar," in *Proceedings of the 19th International Society for Music Information Retrieval Conference* (ISMIR, 2019), 462–469.

27 Steedman, "The Blues and the Abstract Truth: Music and Mental Models."

28 Carl Pollard and Ivan A. Sag, *Head-driven Phrase Structure Grammar* (Chicago: University of Chicago Press, 1994); Stefan Müller, *Head-Driven Phrase Structure Grammar: Eine Einführung* (Tübingen, 2013).

29 Satoshi Tojo, Yoshinori Oka, and Masafumi Nishida, "Analysis of Chord Progression by HPSG," in *Proceedings of the 24th IASTED International Conference on Artificial Intelligence and Applications* (Innsbruck, Austria, 2006), 305–310.

30 Martin Rohrmeier, "A Generative Grammar Approach to Diatonic Harmonic Structure," in *Proceedings of the 4th Sound and Music Computing Conference (SMC)*, ed. Haralambos Spyridis et al. (2007), 97–100; Martin Rohrmeier, "Towards a Generative Syntax of Tonal Harmony," *Journal of Mathematics and Music* 5, no. 1 (2011): 35–53, ISSN: 17459737, doi:10.1080/17459737.2011.573676, http://www.tandfonline.com/doi/abs/10.1080/17459737.2011.573676.

imperfect, evaded, and deceptive cadential strategies.[31] Jonah Katz proposed another syntactic model of Blues (the differences between Steedman's, Katz's and my own model will be discussed in detail in section 5.5 on Blues).[32] Although several Schenkerian concepts such as "interruption" suggest dependency complexity beyond context-free structures, models of higher (mildly) context-sensitive complexity have barely been explored.[33]

Because of the particularly close epistemological connection between generative theory building and computational modeling, numerous approaches go hand in hand with computational implementations.[34] Grammar models have found increasing use in the computational modeling of music and algorithmic composition.[35] This includes particularly the early landmark syntax model by Mark Steedman and its extensions[36], as well as the Generative Syntax Model (GSM)[37]. In related lines of research, there have

31 Martin Rohrmeier and Markus Neuwirth, "Towards a Syntax of the Classical Cadence," in *What is a Cadence? Theoretical and Analytical Perspectives on Cadences in the Classical Repertoire*, ed. Markus Neuwirth and Pieter Bergé (Leuven: Leuven University Press, 2015), 287–338.

32 Jonah Katz, "Harmonic Syntax of the Twelve-Bar Blues Form: A Corpus Study," *Music Perception* 35, no. 2 (2017): 165–192.

33 Rohrmeier and Pearce, "Musical Syntax I: Theoretical Perspectives."

34 Pearce and Rohrmeier, "Musical Syntax II: Empirical Perspectives."

35 Édouard Gilbert and Darrell Conklin, "A Probabilistic Context-free Grammar for Melodic Reduction," in *Proceedings of the International Workshop on Artificial Intelligence and Music, 20th International Joint Conference on Artificial Intelligence (IJCAI-07), Hyderabad, India*, 2007 (2007), 83–94; Donya Quick and Paul Hudak, "Grammar-Based Automated Music Composition in Haskell," in *Proceedings of the First ACM SIGPLAN Workshop on Functional Art, Music, Modeling & Design (FARM '13)* (2013), 59–70, doi:10.1145/2505341.2505345; Samer A. Abdallah and Nicolas E. Gold, "Comparing Models of Symbolic Music Using Probabilistic Grammars and Probabilistic Programming," in *Joint Sound and Music Computing Conference and International Computer Music Conference, Athens, Greece* (2014); Samer Abdallah, Nicolas Gold, and Alan Marsden, "Analysing Symbolic Music with Probabilistic Grammars," in *Computational Music Analysis*, ed. David Meredith (Cham: Springer, 2016), 157–189, doi:10.1007/978-3-319-25931-4_7; Klaus Frieler, "Constructing Jazz Lines: Taxonomy, Vocabulary, Grammar," in *Jazzforschung heute: Themen, Methoden, Perspektiven*, ed. M. Pfleiderer and W.G. Zaddach (Berlin: Edition EMVAS, 2019), 103–132; Tillman Weyde, "Grammatikbasierte harmonische Analyse von Jazzstandards mit Computerunterstützung," in *KlangArt-Kongreß 1995* (1995), 315–324; Tillman Weyde, "Automatic Semantic Annotation of Music with Harmonic Structure," 2007, Dan Tidhar, *A Hierarchical and Deterministic Approach to Music Grammars and Its Application to Unmeasured Preludes* (Dissertation.de, 2005)

36 Steedman, "A Generative Grammar for Jazz Chord Sequences"; Steedman, "The Blues and the Abstract Truth: Music and Mental Models"; Mark Granroth-Wilding and Mark Steedman, "Statistical Parsing for Harmonic Analysis of Jazz Chord Sequences," in *Proceedings of the International Computer Music Conference* (2012), 478–485; Marc Chemillier, "Toward a Formal Study of Jazz Chord Sequences Generated by Steedman's Grammar," *Soft Computing* 8, no. 9 (2004): 617–622; François Pachet, "Computer Analysis of Jazz Chord Sequences: Is Solar a Blues?," in *Readings in Music and Artificial Intelligence*, ed. Eduardo Reck Miranda (2000), 85–113; Mark Granroth-Wilding and Mark Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences," *Journal of New Music Research* 43, no. 4 (2014): 355–374.

37 Bas De Haas et al., "Modeling Harmonic Similarity Using a Generative Grammar of Tonal Harmony," in *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)* (2009); Daniel Harasim, Martin Rohrmeier, and Timothy J O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax," in *Proceedings of the 19th Conference of the International Society for Music Information Retrieval* (ISMIR, 2018), 152–159; Daniel Harasim, Timothy J. O'Donnell, and Martin Rohrmeier, "Harmonic Syntax in Time: Rhythm Improves Grammatical Models of Harmony," in *Proceedings of the 20th Conference of the International Society for Music Information Retrieval* (ISMIR, 2019), 335–342.

been several approaches towards modeling the GTTM[38] as well as some components of Schenkerian analysis using formal and computational approaches.[39]

Apart from the above, another central issue concerns the question of whether the structural principles that are postulated in syntactic approaches to music are domain-specific or derive from overarching domain-general principles of cognition—a question that has similarly been raised about the human capacity of language. While the principle of recursion has been argued to be primary for language[40], there have been other approaches that situate such structural models within our broader cognitive capacities.[41] For instance, Ray Jackendoff has extended the spirit of generative theory building from the linguistic domain to a general understanding of the mind,[42] as exemplified, for instance, by a generative model of coffee making.[43] In a similar vein, Neil Cohn proposes the application of generative trees to the story sequence of comic strips.[44] The overarching idea that recursive hierarchical principles of structuring may be a domain-general mechanism of the human mind and may indeed trace back to complex planning has also been corroborated by the work of Mark Steedman.[45]

38 Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, "Implementing 'A Generative Theory of Tonal Music'," *Journal of New Music Research* 35, no. 4 (2006): 249–277; Masatoshi Hamanaka, Keiji Hirata, and Satoshi Tojo, "Implementing Methods for Analysing Music Based on Lerdahl and Jackendoff's Generative Theory of Tonal Music," in *Computational Music Analysis*, ed. David Meredith (Cham: Springer, 2016), 221–249.

39 Stephen W. Smoliar, "A Computer Aid for Schenkerian Analysis," *Computer Music Journal* 4, no. 2 (1980): 41–59; Alan Marsden, "Generative Structural Representation of Tonal Music," *Journal of New Music Research* 34, no. 4 (2005): 409–428, doi:http://dx.doi.org/10.1080/09298210600578295; Alan Marsden, "Schenkerian Analysis by Computer: A Proof of Concept.," *Journal of New Music Research* 39, no. 3 (2010), doi:10.1080/09298215.2010.503898; Phillip Kirlin and Paul Utgoff, "A Framework for Automated Schenkerian Analysis," in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)* (2008), 363–368; Phillip B. Kirlin and David D. Jensen, "Probabilistic Modeling of Hierarchical Music Analysis," in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)* (2011), 393–398.

40 Marc D. Hauser, Noam Chomsky, and W. Tecumseh Fitch, "The Faculty of Language: What Is It, Who Has It, And How Did It Evolve?," *Science* 298, no. 5598 (2002): 1569–1579.

41 E.g., Maurício Dias Martins and William Tecumseh Fitch, "Investigating Recursion Within a Domain-specific Framework," in *Language and Recursion*, ed. Francis Lowenthal and Laurent Lefebvre (Heidelberg: Springer, 2014), 15–26; Maurício de Jesus Dias Martins et al., "Representing Visual Recursion Does not Require Verbal or Motor Resources," *Cognitive Psychology* 77 (2015): 20–41.

42 Ray Jackendoff, *Language, Consciousness, Culture: Essays on Mental Structure* (Cambridge, MA: The MIT Press, 2007); Ray Jackendoff, *Consciousness and the Computational Mind* (Cambridge, MA: The MIT Press, 1987); Ida Toivonen, Piroska Csuri, and Emile Van Der Zee, *Structures in the Mind: Essays on Language, Music, and Cognition in Honor of Ray Jackendoff* (Cambridge, MA: The MIT Press, 2015).

43 Ray Jackendoff, "Parallels and Nonparallels between Language and Music," *Music Perception* 26, no. 3 (2009): 195–204, doi:10.1525/mp.2009.26.3.195.

44 Neil Cohn, "Visual Narrative Structure," *Cognitive Science* 37, no. 3 (2013): 413–452, doi:10.1111/cogs.12016; Neil Cohn, "Climbing Trees and Seeing Stars: Combinatorial Structure in Comics and Diverse Domains," in *Structures in the Mind: Essays on Language, Music, and Cognition in Honor of Ray Jackendoff*, ed. Ida Toivonen, Piroska Csúri, and Emile van der Zee (Cambridge, MA: The MIT Press, 2015), 379–392.

45 Steedman, "The Blues and the Abstract Truth: Music and Mental Models"; Mark Steedman, "Plans, Affordances, and Combinatory Grammar," *Linguistics and Philosophy* 25, nos. 5–6 (2002): 723–753; Mark Steedman, "The Emergence of Language," *Mind & Language* 32, no. 5 (2017): 579–590.

Investigating the cognitive embedding of the music faculty, the discussion has predominantly focused on the relations between music and language.[46] In the light of such discussions, Aniruddh Patel proposed the hypothesis of shared neural resources between music and language.[47] Complementary neuroscientific evidence supports the view that harmonic structure is processed in a hierarchical fashion in analogy to language.[48] From a theoretical perspective, Jonah Katz and David Pesetsky proposed their influential identity thesis[49] of linguistic and musical syntax. This hypothesis proposes that the abstract formalism governing language and music may be identical while the building blocks are different. This requires musical structure to not only be context-free, but also exhibit mild context-sensitivity[50] (i.e. requiring also the application of internal merge or the movement operation in language, or a similar mechanism[51]). The discussion around the identity thesis is picked up at the end of this article. In order to advance this debate, however, further music-theoretical research is required to gather more evidence about complex dependency structures in music.

The purpose of this article is to propose a generative syntax of chord sequences in Jazz, integrating insights from the theory of formal languages, computation, linguistics, as well as *Neo-Riemannian* and *Tonfeld* theory, and to formulate a model that lends itself to computational implementation and exploration. This paper consists of two main parts. The first part outlines a model of harmonic syntax that proposes tree-based syntactic analyses based on diatonic derivations of harmonic phenomena. The second outlines how the main aspects of phrase structure and musical form in Jazz are expressed in the syntactic core. An independent extension of the formalism will be formulated in a second article: starting from the observation that certain harmonic relations cannot be easily

---

46 Patrick Rebuschat et al., *Language and Music as Cognitive Systems* (Oxford: Oxford University Press, 2012); Rie Asano and Cedric Boeckx, "Syntax in Language and Music: What is the Right Level of Comparison?," *Frontiers in Psychology* 6 (2015): 942; Martin Rohrmeier et al., "Principles of Structure Building in Music, Language and Animal Song," *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, no. 1664 (2015): 20140097; Jackendoff and Lerdahl, "The Capacity for Music: What is It, and What's Special About It?"; Michael A. Arbib, ed., *Language, Music, and the Brain: A Mysterious Relationship* (Cambridge, MA: MIT Press, 2013).

47 Aniruddh D. Patel, *Music, Language, and the Brain* (Oxford: Oxford University Press, 2008); Aniruddh D. Patel, "Language, Music, and the Brain: A Resource-sharing Framework," in *Language and Music as Cognitive Systems*, ed. Patrick Rebuschat et al. (2012), 204–223.

48 Aniruddh D. Patel et al., "Processing Syntactic Relations in Language and Music: An Event-Related Potential Study," *Journal of Cognitive Neuroscience* 10, no. 6 (1998): 717–733; Stefan Koelsch et al., "Interaction between Syntax Processing in Language and in Music: An ERP Study," *Journal of Cognitive Neuroscience* 17, no. 10 (2005): 1565–1577; Stefan Koelsch et al., "Processing of Hierarchical Syntactic Structure in Music," *Proceedings of the National Academy of Sciences* 110, no. 38 (2013): 15443–15448.

49 Jonah Katz and David Pesetsky, "The Identity Thesis for Language and Music" (2011), http://ling.auf.net/lingbuzz/000959.

50 David Jeremy Weir, "Characterizing Mildly Context-Sensitive Grammar Formalisms" (PhD diss., University of Pennsylvania, 1988).

51 Stefan Müller, *Grammatiktheorie* (Tübingen: Stauffenburg, 2010); Adger, *Core Syntax: A Minimalist Approach*.

explained using diatonic derivations, it argues that complex harmonic phenomena in Jazz can be explained through a model of extended tonality based on *Neo-Riemannian* and *Tonfeld* theory.[52]

2. MOTIVATION

Jazz harmony may be regarded as part of an extended common practice—one that continues the tradition of Classical eighteenth- and nineteenth-century harmony—since otherwise it could not be intuitively understood by (enculturated) listeners. The system of Jazz harmony, in addition, constitutes an established language that reaches into numerous styles of contemporary tonal music—similar to the language of the Classical style pervading a large part of the common practice.[53] Drawing on the canonical repertoire of Jazz, this article proposes a theory of its harmonic syntax that models most of the tonal language that has been disseminated in the "Great American Song Book" and the different "Real Books", and spans a broad range of traditional and extended tonal styles.The theory models chord sequences as they appear in lead sheets, which are commonly used by musicians as a foundation for improvisation and the transmission of pieces. Other aspects, such as complex chord forms and upper structures, the analysis of solo transcriptions, schema theory or reductive voice-leading analyses are not the focus of this text.Such directions have been taken in the work by Steve Larson, Henry Martin, David Heyer, Dariusz Terefenko, and others, as mentioned above.

The following sections lay the foundations for the syntactic approach in terms of reflections on the relation of structure and interpretation, kinds of harmony, and hierarchicality. The proposed syntactic theory is outlined in section 3.

2.1. *Structure and Interpretation, and Rethinking Grammaticality*

Fundamentally, musical hearing and analysis are "hearing as" and "reading as".[54] Functional tonal harmony is thus at its heart distinguished by the fact that it is closely linked with *interpretation* and *interpretability*, and this sets it apart from a mere characterization

---

52 Richard Cohn, *Audacious Euphony: Chromatic Harmony and the Triad's Second Nature* (New York: Oxford University Press, 2012); Edward Gollin and Alexander Rehding, *The Oxford Handbook of Neo-Riemannian Music Theories* (Oxford: Oxford University Press, 2011); Bernhard Haas, *Die neue Tonalität von Schubert bis Webern: Hören und Analysieren nach Albert Simon* (Wilhelmshaven: F. Noetzel, Heinrichshofen-Bücher, 2004).

53 William E. Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven* (New York: Oxford University Press, 2001).

54 See also Joseph Dubiel, "Music Analysis and Kinds of Hearing-As," *Music Theory and Analysis* 4, no. 2 (2017): 233–242; Ludwig Wittgenstein, *Philosophical Investigations*, ed. G.E.M. Anscombe, G.H. von Wright, and Rush Rhees (London: Macmillan, 1953).

based on plain (surface) chord successions. Its underlying interpretation differentiates functional harmony from vertical voice-leading sonorities (which occur because they result from multiple simultaneous lines, rather than being chords in their own right) and modal uses of chords (as in pre-Baroque or some forms of Rock music[55]; see also the next subsection). Here, *interpretation* involves to ascribe a secondary structure to a given chord, such as its tonal role and its function within the hierarchical network of relations. Differences in the interpretation may be expressed in different analyses of the same surface chord (for instance, *I* vs. *V/IV*).[56] Since functional analyses like prolongational or preparatory relations may be nested, the resulting overarching interpretation may express a whole set of hierarchical fine-grained differentiations in the corresponding hearing; these, together with a cognitive mechanism of in-time processing[57], may characterize the cognitive correlates of some higher-order listening experiences, such as the build-up of closure, ambiguity, surprise and revision, or build-up and release of musical tension.

Overall, this entails a different perspective on syntax and syntactic modeling: The essence of the syntactic framework is not normative and not to distinguish grammaticality from ungrammaticality, but to express the relation between syntactic structure and its corresponding interpretation that the music affords ("hearing as").[58] Based on this background, it is useful to reflect on different kinds of sonorities and chord use.

2.2. *Kinds of Harmony*

Based on the reasoning above, it is useful to distinguish between different forms of harmony or vertical sonorities for modeling harmonic structure (Figure 1).[59] One main distinction concerns the difference between *functional* and *improper* harmony.[60] Functional harmony refers to a system whose categories fulfill functions in music as structures in their own right (such as implication or closure-building); that is, functional harmony affords for specific kinds of *interpretation*. Improper harmony ("uneigentliche

55 David Temperley, *The Musical Language of Rock* (New York: Oxford University Press, 2018).

56 Michael Polth, for instance, characterizes a subtle distinction between a chord on scale degree $\hat{5}$ and the dominant (Michael Polth, "Ist die Funktionstheorie eine Theorie von Funktionalität?," *Musiktheorie* 16, no. 4 (2001): 319–324.).

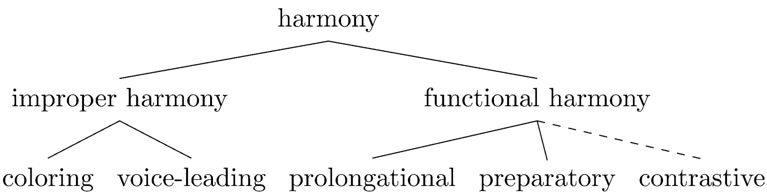
57 Ray Jackendoff, "Musical Parsing and Musical Affect," *Music Perception* 9, no. 2 (1991): 199–229.

58 See the discussion on the foundations of musical syntax in Rohrmeier and Pearce, "Musical Syntax I: Theoretical Perspectives."

59 Salzer makes similar distinctions Felix Salzer, *Structural Hearing: Tonal Coherence in Music* (New York: Dover Publications, 1952).

60 "Improper" is used here not in a pejorative sense, but in analogy to the mathematical term (*improper / uneigentlich*), as in "improper fraction."

Figure 1: The uses and functions of harmony.



Harmonik") by contrast, characterizes occurrences of vertical sonorities that do not act as such building blocks.[61]

2.2.1. Improper Harmony: Coloring and Voice-Leading Sonorities

Not all chords—or better, vertical sonorities—occurring in pieces have their logical origin in functional relations. Improper harmony can arise from vertical co-occurrences in voice-leading contexts that are not "meant" as real chords (i.e., they do not express functional relations and have no such interpretation). The term *coloring harmony* is applied to phenomena in which sonorities are used in ways that abandon tonal gravitation and yield timbre-like effects without functional interpretation.[62] A well-known example of coloring harmony in terms of mixture sonorities is found in Claude Debussy's prelude "La cathédrale engloutie" (Figure 2), where the chords become mere blocks of color and drop their tonal gravitation—despite parts where roots descend in fifths—because of their strictly parallel voice leading over a pedal.[63] Many powerful instances of coloring harmony may be found in the unusual chord progressions in compositions by Wayne Shorter.[64]

Examples of non-functional voice-leading sonorities are found in, for instance, parallel (e.g., fauxbourdon) or passing motions, such as $I/\hat{3}$–(*iii*)–*IV* (–*V*),[65] where the pseudo-chord *iii* results from a leading-tone exchange (Figure 2, right part). Another example is found in the middle section of Thelonious Monk's "Well, you needn't" (see Figure 3). While being framed by a tonal context $F\ G^7 \ldots G^7\ C^7$, the chords in the middle are simply linearly

61 Note that the distinction does not imply that voice leading and harmony are strictly separate, see e.g. Aldwell and Schachter, *Harmony and Voice Leading*.

62 As will be argued in the follow-up article, the application of theories of extended tonality may drastically reduce the need to resort to coloring harmony.

63 See Michael Polth's analysis of non-functionality of this kind, and his argument for the interaction between harmony, voice-leading, rhythm, and texture to induce functionality: Michael Polth, "Nicht System—nicht Resultat: Zur Bestimmung von harmonischer Tonalität," *Musik und Ästhetik* 5, no. 2 (2001): 12–36.

64 See also Steven Strunk, "Notes on Harmony in Wayne Shorter's Compositions, 1964–67," *Journal of Music Theory* 49, no. 2 (2005): 301–332.

65 Note that for reasons of consistency, the chord inversion in this example is not notated using figured-bass style as in Roman-numeral analysis, which would be I^6–*iii*–*IV*(–*V*).

Figure 2: Examples of two cases of improper harmony. The left excerpt is an illustration of the *coloring* use of harmony, from Claude Debussy's Prelude "La cathédrale engloutie", *Preludes, 1^{er} livre, X*, mm. 28–31. The right example is an example of a sonority resulting from voice leading, where the non-functional "chord" "*iii*" results from mere passing motion in the soprano line.

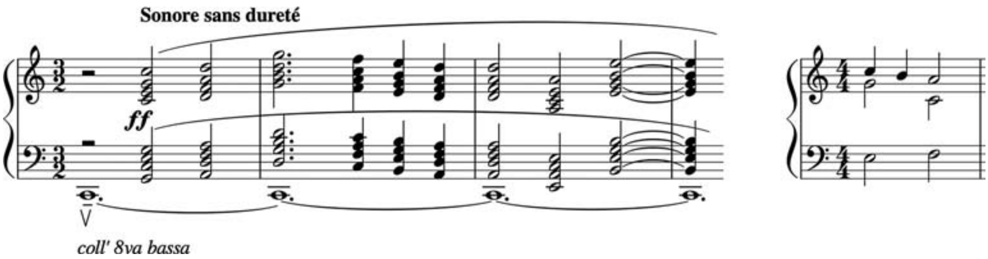


Figure 3: The middle section of Thelonious Monk's "Well, you needn't". The harmonic sequence illustrates a chord sequence that results from non-functional chromatic chord shifts (Rückungen).

chromatically shifted up to B^7 and then down again (an example of a harmonic shift ("Rückung")). A similar case might be attested to the characteristic chromatic ascending riff of Duke Ellington's "Don't get around much anymore."

Notably, such cases of *Rückungen* often appear in chromatically or diatonically ascending (or descending) chord sequences linking two stable chords that provide an outer frame. Further examples are linear diatonic shifts such as *I–ii–iii–ii* (*I*), such as in the middle section of "My funny Valentine."

2.2.2. Functional Harmony

Functional harmony, in turn, may fulfill three main functions: (1) it can *prolong* a harmony (such as *I* or *V*), (2) it can *prepare* a harmony, or (3) it can establish a *contrast* between two harmonies (see Figure 1). These functions are explained in detail below when the core model is outlined. While the assumption of a *contrast* function might arguably not be necessary in traditional diatonic harmony, it may be usefully postulated in the context of extended tonality (as will be expanded on in a follow-up article).

2.3. *Hierarchicality*

The structures resulting from functional harmony operate hierarchically. For instance, in the 1984 edition of their harmony textbook, Kostka and Payne introduce different "levels of harmony."[66] An example may be a whole 4-bar group acting as a prolongation of *V*, or other simple instances of this are already found in discussions concerning the level of granularity of harmonic analyses (e.g., is this bar better analyzed as an instance of *V ii V* or simply *V*?), which point at different levels of harmonic reduction.

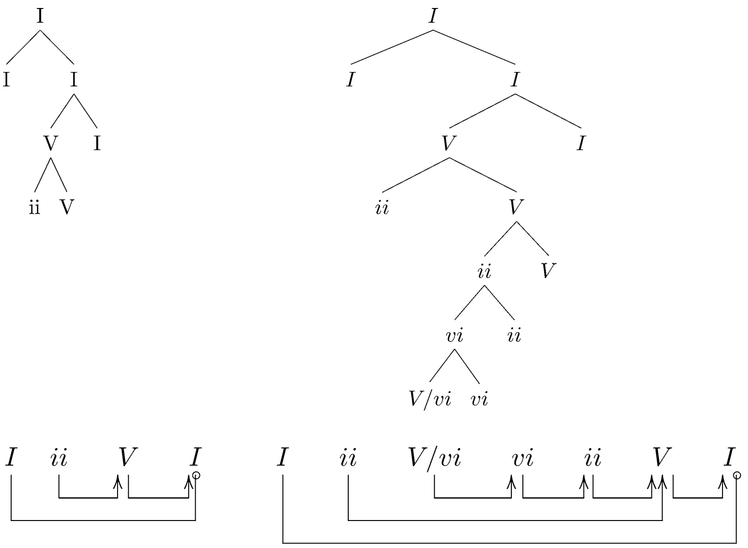
Altogether, hierarchicality in music involves several aspects: (a) Certain chords are more stable than others; for instance, prolongations establish event stability, and goals are more stable than their preparations. Nested dependency structures induce a hierarchy of stability. (b) Harmonic relationships operate at different time-scales and levels of reduction ranging from local to larger segments and sections. (c) Harmonic relations and dependencies may not necessarily be established with the subsequent chord, but between chords that are further apart (because other chords may be inserted in between in terms of elaborations). (d) A chord may govern an entire group of chords (constituent) that have the effect of prolonging their head (for instance, a group of chords prolonging *V*). Heads may stand for the chords they govern in a reduction.

Such distinctions and hierarchical harmonic relations invoke another important distinction, namely whether the goal of the modeling endeavor is merely to reproduce the surface sequence of chords, or whether the model should reproduce the surface sequence of chords along with the relationships that the theorist assumes between them. The former is referred to as *weak* and the latter as *strong generativity*. Strong generativity, which is the purpose of this paper's model, involves capturing the kinds of harmonic functions, hierarchical functional dependencies and nested constituents to model the theorist's *interpretation*. Formal grammars are the tool of choice to express such relationships.

Based on the two sequences *I–ii–V–I* and *I–ii–V/vi–vi–ii–V–I*, figure 4 illustrates the set of hierarchical relations that may be stipulated between chords using a dependency graph and a tree analysis. It exemplifies the hierarchy of stability (top to bottom of the tree), local and non-local harmonic dependencies (e.g., the overarching tonic prolongation or the two *ii–V* preparations in the second graph), different hierarchical timescales (e.g., an overarching *I–V–I* prolongation), and chords governing constituents (e.g., the left being an

66 Kostka and Payne, *Tonal Harmony with an Introduction to 20th-Century Music*. See also Steven Strunk, "The Harmony of Early Bop—A Layered Approach," *Journal of Jazz Studies* 6, no. 1 (1979): 4–53.

Figure 4: Expressing syntactic relations by a dependency graph and a syntax tree.



example of an entire tonic constituent). For music theorists who may not be familiar with the concepts from tree analyses and formal grammars, a brief introduction is provided in the appendix.

3. MODELING DIATONIC HARMONIC RELATIONS

Building on and generalizing previous work[67], this paper proposes an overarching generative syntactic framework for Jazz harmony. Since the canonical Jazz language extends the common practice, many of the diatonic harmonic relations and rules for modulation are similar or identical to common-practice classical harmony.

3.1. *The Core Formalism*

The diatonic syntactic model is grounded in two core rule types for structural dependencies: *prolongation*[68] and *preparation* (see Figure 5). An orthogonal type of rule is harmonic *substitution*, which results in a unary (non-branching) rewrite rule, because it does not connect two chords in terms of a sequential dependency relation.

---

67 Rohrmeier, "A Generative Grammar Approach to Diatonic Harmonic Structure"; Rohrmeier, "Towards a Generative Syntax of Tonal Harmony"; Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence."

68 Note that in this framework the notion of "prolongation" is not used in the wide sense in terms of general syntactic elaboration, as, for instance, in the "prolongational reduction" in the GTTM or, in the Schenkerian sense. Here, "prolongation" is used in a narrow sense indicating the *extension* of a certain harmonic context, such as *I* or *V*.

One core assumption in the grammar is that harmonic sequences are primarily generated within the diatony of the respective major or minor key, and that all non-diatonic chords have to be licensed through certain derivations in order to be able to appear in the context of a diatonic sequence.[69]

Many rules are formulated in a form that generalizes over different chord types, such as $X \rightarrow V/X\ X$. Hence, this formalism is, strictly speaking, not a context-free grammar any more, but an *abstract context-free grammar*.[70] First the grammar and core rules are introduced. Detailed analytial examples are found in section 4.

3.1.1. Symbols

Surface chords are represented in their absolute form (e.g., as Cm^7, F^{maj7}, $F^{\triangle}$ etc.). Interpreted chords at the deeper level are represented in terms of Roman numerals, such as *i*, V^7, *IV*, *V/V*, etc. The applied function operator "/" will be used in the generalized fully recursive fashion, as, for instance, in *V*, *V/V*, *V/V/V*, or *ii/V/V*, etc. Throughout this article, Jazz notation[71] will be used for all chords in surface or scale-degree notation. Accordingly, superscript additions to Roman numerals do *not* express chord inversion, but added or modified chord extensions; if necessary, inversion is notated by specifying the bass note (absolute or relative) with a slash.[72] Augmented chords will be notated with the "+" symbol, as in G^+, half-diminished chords as in $Em^{7\flat5}$ or $Em^{\varnothing}$, and fully diminished seventh chords as in D^0.

When dealing with key changes in the context of modulations or tonicizations, Roman numerals will be used to express local subordinated keys, for instance *V*, *V/V*, *ii*, etc. Since keys will be notated in terms of features of chords, there will be no ambiguity whether a symbol represents a key or a chord. Absolute keys are notated with symbols such as *A(maj)* or *Cm*.

3.1.2. Prolongation as Coordination

Prolongation (or coordination) expresses the core relation that two elements of the same type can be coordinated instantiating an extension of a single higher-order element. This

---

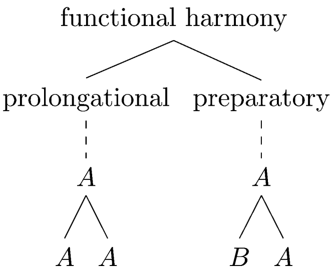
69 Even though it is not a core topic of this article, the grammar assumes further that Jazz pieces in other modes mostly follow the logic of the closest major / minor mode under appropriate chord changes, i.e. Dorian and Phrygian behaving similar to minor (Aeolian); Mixolydian, Lydian behaving similar to major (Ionian); and, Locrian, to occur very rarely as the main mode of a piece.

70 Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax."

71 See, e.g., Levine, *The Jazz Theory Book*; Terefenko, *Jazz Theory: From Basic to Advanced Study*.

72 David Temperley provides an interesting analysis of the comparably rare occurrence of chord inversions in Jazz based on communicative pressure, see David Temperley, "Communicative Pressure and the Evolution of Musical Styles," *Music Perception: An Interdisciplinary Journal* 21, no. 3 (2004): 313–337.

Figure 5: The syntactic forms of the main syntactic dependencies in diatonic functional harmony: *prolongation* and *preparation*.



makes it possible to express that tonic elements are prolonged over larger time spans, while themselves being prepared with other chords, or that several dominants are coordinated as a genuine constituent that may prepare a tonic element. Coordination is expressed by a simple abstract generative rule (rule 1) that generalizes over the element type it applies to.[73] The analyses in Figures 4 and 23 provide examples of the prolongation of the tonic *I*.

$$X \rightarrow X\,X \qquad (1)$$

Lerdahl and Jackendoff draw a further essential distinction between weak and strong prolongation.[74] Such a distinction may be incorporated in the present formalism as well such that for a different inversion or voicing X' of the same chord X, one of the two chords may be distinguished as the head if required. The corresponding pair of rules modeling weak prolongation would be:

$$X \rightarrow X'\,X \qquad (2)$$
$$X \rightarrow X\,X' \qquad (3)$$

3.1.3. Preparation

The second core dependency relation is *preparation*, and it exhibits the following abstract form:

$$X \rightarrow Prep/X\;X \qquad (4)$$

Here, *Prep*/*X* summarizes over the set of all chords that can prepare *X*. In the concrete context of the diatonic framework, there are several options for instantiating such preparations (for instance, in terms of (diatonic) descending fifths progressions, *IV*–

73 See also Granroth-Wilding and Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences", for an account of harmonic coordination in a functional way.

74 Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*.

V, *(♭)VI–V* or *vii°–I* progressions). Also, non-diatonic applied dominants can prepare a given chord. The following list provides a selection of cases that instantiate rule (4):

$$X \rightarrow \Delta/X\ X \qquad (5)$$

$$V \rightarrow IV\ V \qquad (6)$$

$$V \rightarrow \flat VI\ V \qquad (7)$$

$$X \rightarrow vii^{\circ}/X\ X \qquad (8)$$

$$X \rightarrow V/X\ X \qquad (9)$$

In rule (5) above, Δ refers to the diatonic fifth (such as in *ii–V–I*, or *Dm G C* in *C* major), while *V/X* in rule (9) refers to the applied (non-diatonic) dominant chord (such as in *V/V–V–I* or D^7*–G–C* in *C* major). This set of rules not only models descending diatonic or applied fifths sequences, but also enables one to derive typical Jazz cadences.[75] Note that the rules presented here are based on the scale degree and type (major, minor, half-diminished, etc.), but abstract from chord additions (such as 7, ♭9, etc.). In analysis, however, taking into account typing constraints over chord additions may help to disambiguate and infer likely chord analyses (e.g. $Xmaj^7$ as an instance of a dominant chord, or $Xm^{7\flat5}$ as a chord on the second scale degree in minor). Such typing constraints aid the (probabilistic) disambiguation of the inference of scale degree and key during parsing.

Note that direct *IV–V* progressions in Jazz are rare compared with *ii–V* relations. However, at a deeper level the *IV* is still very present, for instance as a tonal center for the path to the *V* in middle sections (see, as an example, Duke Ellington's "Take the A-Train," Figure 9). This reveals that secondary dominants have some very immediate surface-related relevance, whereas subdominants still operate on larger (deeper) timescales.

In addition to the use of the diminished chord as a dominant substitute, it is useful to assume further rules to model its use in voice-leading contexts, in particular, the auxiliary diminished and the descending chromatic approach, as outlined by the rules below.[76] Note that because Jazz chords and corresponding rules often do not incorporate or spell out inversion, progressions like *I°–I* or *♯IV°–I/*$\hat{5}$ are covered by the same rule. The second rule models descending chromatic voice-leading preparations like *♭vi°–V* or $ii^{7\flat5}$*–i*, which are relatively rare (and because of their voice-leading based origin, they may also—like the others—be considered "improper harmony").

75 The derivation of the half cadence, and whether it constitutes a context-free or context-sensitive phenomenon, shall not be discussed in this article. Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence", provide a detailed perspective on the analysis of the half cadence.

76 Following the classification of diminished chords in Jazz by Mulholland and Hojnacki, *The Berklee Book of Jazz Harmony*.

$$X \rightarrow X^{o}\ X \tag{10}$$

$$X \rightarrow \flat ii^{o}/X\ X \tag{11}$$

$$X \rightarrow ii^{7\flat 5}/X\ X \tag{12}$$

Another rule that may be useful to assume is the plagal dependency (rule 13). Because Jazz harmony exhibits a strong preference towards authentic preparatory progressions, the plagal rule occurs much more rarely in Jazz[77] than in, say Blues or Rock. One example of a phrase-initial plagal progression (*iv–I*) is found at the beginning of Cole Porter's "All of you." Although the plagal "relaxation" is in its musical effect fundamentally different from the strong implicative power of the authentic step of (applied) dominants, the derivational form of the plagal progression still matches the left-branching form of a preparation.[78] This entails that the mere derivation of a left branch from a goal event is not sufficient to define a dominant.[79] The main application of the plagal rule lies in the harmony of the Blues form (see section 5.5).

$$X \rightarrow IV/X\ X \tag{13}$$

3.1.4. Substitution

Another way to derive complex chords is by substitution.[80] The most frequent case in Jazz is the *tritone substitution*. Substitution may also play a role at a deeper level, such as in *♭III* being introduced as a substitution of *i* in minor, which may establish the seed of a modulation to the key of *♭III*. The generalized form of the unary substitution rule is:

$$X \rightarrow Sub/X \tag{14}$$

Sub/X summarizes over substitutions of *X*. For example, *♭III* may instantiate this function as a *Sub/I* in $I \rightarrow \flat III/I$. As a more complex example, the derivation of the middle section of "Night and Day" requires the pendulum *E♭–C–E♭–C*, which can only be derived within the diatonic framework through a tonic prolongation (of the overarching *C* major tonic *I*)

77 Daniel Shanahan and Yuri Broze, "A Diachronic Analysis of Harmonic Schemata in Jazz," in *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music* (2012), 909–917.

78 The assumption of the plagal rule creates an ambiguity in a sequence like *IV–V–I*, where *IV* now can be analyzed as a left branch of *I* (which provides another argument for the sparse use of the plagal rule outside of domains like the Blues). Ambiguities of this kind can be addressed with probabilistic modeling, Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax."

79 In contrast to an argument along those lines by Fred Lerdahl, *Tonal Pitch Space* (New York: Oxford University Press, 2001), 214ff.

80 For an exploration of variants of the *ii–V–I* sequence, see, for instance, Dariusz Terefenko, "Jazz Transformations of the ii7–V7–I Progression," *Current Research in Jazz* 1 (2009).

that features substitutions to the relative *E♭* major (*♭III*) of the parallel *C* minor key (*i*). The explanation of this middle section will be simplified through the notion of octatonic equivalence as proposed in the extended tonality framework in the follow-up article. Two central cases of harmonic substitutions in Jazz are the tritone substitution and the "backdoor dominant" *♭VII*, both of which are discussed in detail in the next section.[81]

3.2. *Deriving Out-Of-Key Chords and Modulation*

The diatonic framework outlined above only models dependencies within the key and its diatony as well as applied dominants. The occurrence of out-of-key chords has to be licensed through specific rules that embed them in a diatonic framework. There are at least four ways by which non-diatonic harmony can be integrated into a diatonic framework: *applied dominants*, *modulation*, the *tritone substitution*, and *borrowing* from the parallel key (*modal interchange*). Applied dominants were already described above, and the others are characterized below.

For the modeling of modulation, tonicization and other phenomena involving key change or out-of-key chords, a *key* feature is assumed. Every nonterminal category has a *key* feature attached that expresses its key context. For instance, $I_{key=A\flat}$ denotes a tonic in *A♭* major and $V/ii_{key=Dm}$ denotes an applied dominant to the *ii* in the key of *D* minor. The key feature can also be denoted relative terms, such as $I_{key=\flat III}$, denoting the tonic in the key of *♭III*, or $ii_{key=V/V}$, denoting the supertonic in the key of the dominant of the dominant (key). This relative key referencing is of particular importance for modulation and recursively embedded key relations.

3.2.1. Tonicization and Modulation

Modulation occurs when a diatonic scale degree (other than the current governing tonic) in a certain key context becomes the new tonic *I* opening its own key context for subsequent derivations.[82] From the perspective of the formalism, there is no formal difference between

81 Generally, the introduction of a unary rule such as substitution does not affect the overall properties of the grammar. A grammar with unary rewrite rules can be rewritten into a weakly equivalent grammar without unary rewrite rules that generates the same set of sequence (Schöning, *Theoretische Informatik—Kurz Gefasst*; Sipser, *Introduction to the Theory of Computation*). However, an implementation of unary substitutions needs to take potential cycles of substitutions into account.

82 Here, the proposed model differs from the one by Mark Granroth-Wilding and Mark Steedman, who do not distinguish between diatonic and non-diatonic derivations and do not assume constraints on modulation in their grammar in terms of their development rule (Granroth-Wilding and Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences"). In essence, they only model chord roots and their relative steps, which results in an efficient, concise formalism. However, a radically relative approach like this may not account for the derivation restrictions or specific preferences that come within the diatonic, affecting distinctions in probability distributions inside and outside the diatonic.

modulation and tonicization because this difference rather concerns the extent of the modulating segment rather than a principal difference. The change of tonal center in the context of a modulation can be expressed as follows:

$$X_{key=Y} \rightarrow I_{key=X/Y}, \qquad X \neq I \tag{15}$$

Accordingly, modulating constitutes a form of casting, in which the scale degree of the chord symbol is functionally applied to the superordinate key in the key feature. Rule (16) expresses the change of mode from major to minor and vice versa (using a function *inv* that inverts the mode).

$$I_{key=Y} \rightarrow I_{key=inv(Y)} \tag{16}$$

Importantly, there are also tonicization phenomena in which the key context is not manipulated by introducing a new *I*, but by having a (relative) *V* become the new head of a subtree with its own key scope. This is necessary, for example, when a remote *V*, such as E^7 in the overarching context of a sequence in the key of *C*, establishes its own diatonic context (in *A* major); in this case, a new diatonic context, like *ii (= Bm)* requires the key feature of *A*, which has to be established by the *V* if the tonic *A* is not present. The following rule provides the formalization of this phenomenon.

$$V/X_{key=Y} \rightarrow V_{key=X/Y} \tag{17}$$

Instances of both kinds of tonicization are found in several examples in this article, for instance in the analyses of "Satin Doll" (Figure 7) and "Blues for Alice" (Figure 20). An even more general form of this rule is $X/Y_{key=Z} \rightarrow X_{key=Y/Z}$, which applies to types other than the dominant. It is assumed, however, that this generality may not be required.

3.2.2. Dominant Substitution

Another way to introduce non-diatonic chords into a diatonic frame is the dominant substitution, and in particular, the tritone substitution, which constitutes a very common speciality of the Jazz language.

For any dominant seventh chord, the equivalence of the tritone between the third and the seventh licenses the exchange of the root of the surface chord by a tritone.[83] This substitution is formally expressed by a unary substitution rule (18). However, this rule

83 For a subtle analysis differentiating the tritone substitution and augmented sixth chords based features such as their voice leading characteristics and harmonic function, see Nicole Biamonte, "Augmented-Sixth Chords vs. Tritone Substitutes," *Music Theory Online* 14, no. 2 (2008).

avoids to exchange the root V^7 for $\flat II^7$. The latter would have no functional *interpretation* (!) since $\flat II$ has no role in the diatony and could not (for example) be further diatonically elaborated. Thus, by exchanging the key feature by a tritone, which maintains consistency and retains the chord as an instance of the dominant type, the formalization also allows for the common case of further recursive derivations in the new key, e.g. $V/V_{key=I}$–$ii_{key=\flat v}$–$V_{key=\flat v}$–I (as in the analysis of "Satin Doll", Figure 7).[84]

$$V^{(7)}/X_{key=Y} \longrightarrow V^{(7)}{}_{key=\flat V/X/Y} \qquad (18)$$

Because of its derivation through the characteristic tritone of the dominant-seventh chord, it remains questionable whether the tritone substitution can be fully equivalently applied to tonic or subdominant chords.[85]

One typical example of the tritone substitution of a dominant is found in the main motif of Dizzy Gillespie's "A Night in Tunisia" (see Figure 6). The tritone substitution often triggers further recursive elaboration or modulation, such as in the derivation of *ii* preparing *V* in the new key domain, or deriving larger sequences of chords that are remote from the original diatonic. At a large scale, the tritone substitution can be be more abstractly applied to govern a structure at a deep level. For instance, it can constitute the head of the middle section (substituting for *V*). A piece like Miles Davis's "So What" (see Figure 15) illustrates this.

The occurrence of a chord with the surface of $\flat II$ creates a potential ambiguity between the tritone substitution and the Neapolitan predominant. However, the use of a Neapolitan chord is, arguably, rare (if not entirely absent) in Jazz, which constitutes a major difference to Classical common practice harmony. Occurrences of $\flat II$ chords or modulations to $\flat II$ predominantly occur as instances of a tritone substitution of the dominant.[86] The standard "Blue Bossa" may serve as a classic example, in which the second part modulates instead to the dominant to its tritone substitution $V^7/\flat V$ before returning to the tonic key—while an interpretation in terms of a Neapolitan would be nonsensical here.

Another case of a dominant substitution in Jazz is the so-called "backdoor dominant," the progression $\flat VII$–I. In analogy to the rule for the tritone substitution, it may be expressed as a change of the underlying keyframe. In this case, the backdoor dominant

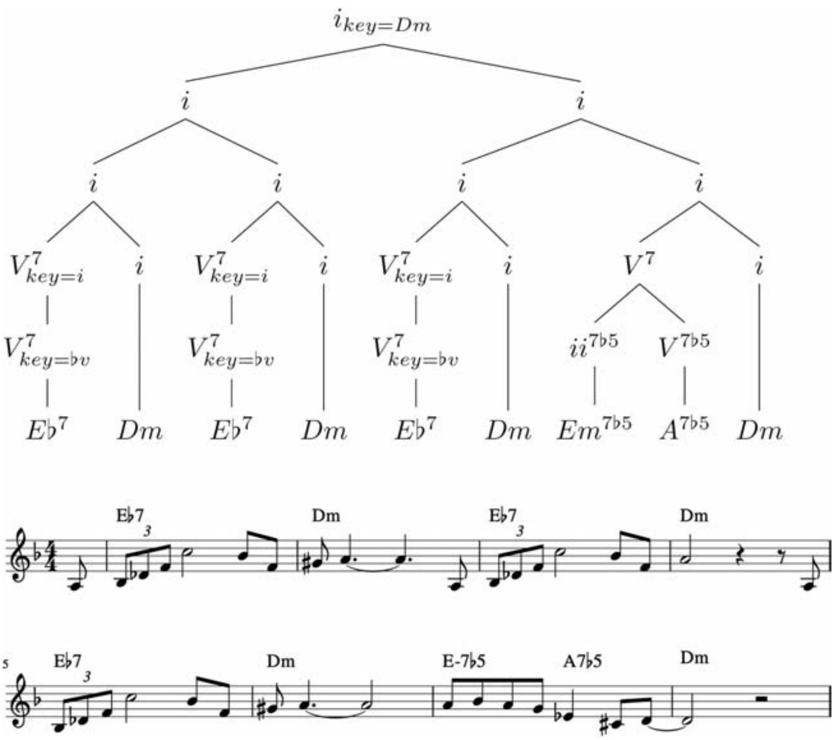
---

84 See also Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax," for computational implementation of this understanding of the tritone substitution.

85 This is valid only under the assumption that subdominants are not generally modeled as kinds of secondary dominants; the discussion whether tritone substitution may also be extended to subdominant or tonic harmonies is encompassed and reframed by the octatonic framework of extended Jazz tonality.

86 Nonetheless it would be possible to incorporate a rule such as $V \rightarrow \flat II_{/3}\ V$ (or as $iv^{\natural 5\flat 6}$) to model the Neapolitan predominant (in first inversion).

Figure 6: The initial motif from Dizzy Gillespie's "A night in Tunisia" illustrates the tritone substitution as formalized in rule (18).



is generated from the dominant of the relative (♭*III*) of the minor tonic. This makes it possible that the backdoor dominant may recursively generate its own preparations from the key of ♭*III*. It may therefore be modeled with the following rule:

$$V^{(7)}/X_{key=Y} \longrightarrow V^{(7)}{}_{key=\flat III/X/Y} \qquad (19)$$

3.2.3. Borrowing

Borrowing results when a diatonic chord from the parallel minor/major key replaces the chord of the same function in the original key. The unrestricted formalization of borrowing can be formulated through inverting the mode of the key feature of the chord symbol:

$$X_{key=Y} \longrightarrow X_{key=inv(Y)} \qquad (20)$$

This rule, in fact, constitutes a generalization of the mode changing rule (16). Note that this rule only applies when the scale degree of the chord root is shared in both modes (e.g. *iii* in major could not undergo borrowing in minor).

4. DISCUSSION

The explanatory power of the two principles of preparation, in particular dominant preparation, and prolongation combined in generative derivation is remarkable. A large number—maybe even the majority—of classical Jazz standards can be analyzed mainly through fifth relations, their elaborations and substitutions in recursive application.

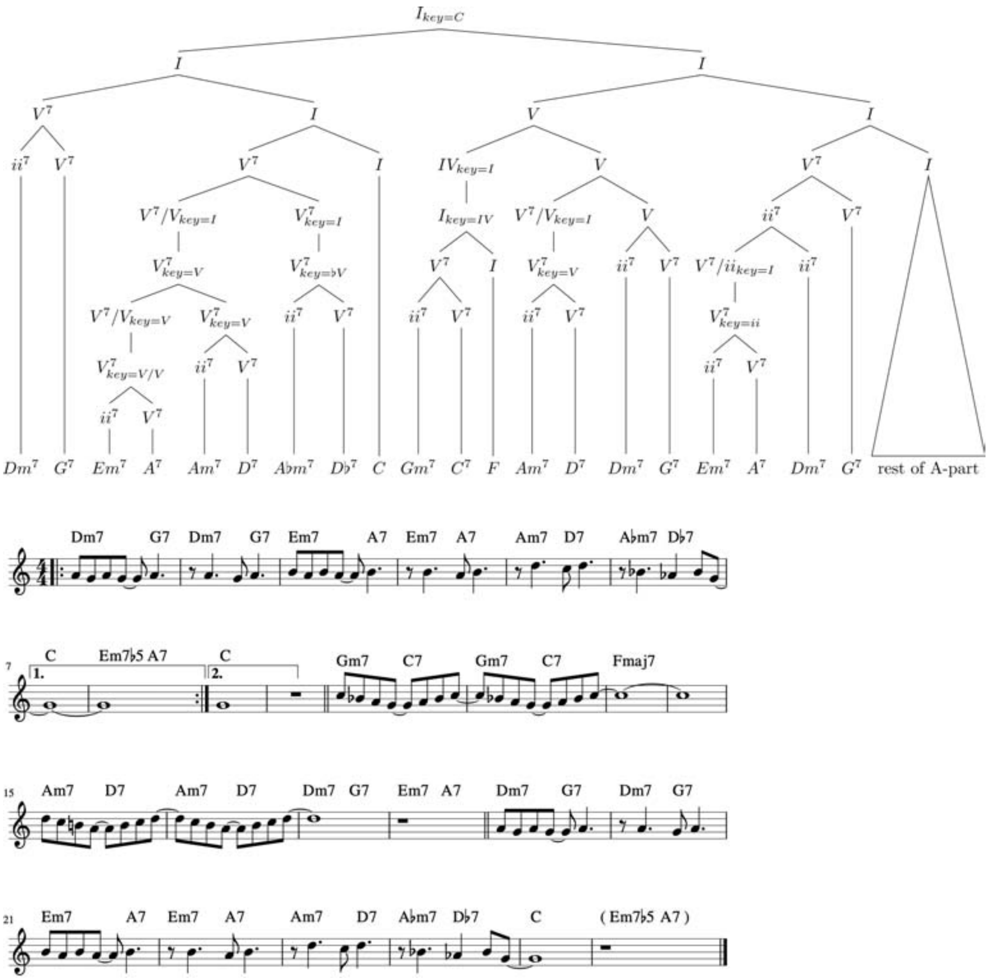
4.1. *Analytical Examples*

A basic example is given by the analysis of Duke Ellington's "Take the A-Train" (Figure 9). The example illustrates that the A part can by generated from a simple overarching *I–V–I* tonic prolongation (similarly to the example in Figure 4), where the dominant is prepared twice, by *ii* and its applied dominant *V/V*. The middle section is derived from an overarching *V* constituent, where *V* is elaborated by three distinct preparations with *IV*, *V/V* and *ii*. The top level of the tree connects the first part, the middle section and the repetition of the first part in an overarching *I–V–I* dependency.

A fully worked more complex example is found in Duke Ellington's "Satin Doll" (Figure 7). The first part (mm. 1–8) constitutes an overarching tonic constituent, which is elaborated by two preparatory constituents derived from the dominant. The first dominant constituent results in the initial *ii–V* sequence that sets up a nonlocal dependency to the phrase-final tonic. The second dominant constituent first features a recursive preparation of the dominant: the dominant *V* is prepared by an applied *ii–V* in its own key (of *V*, i.e. *G* major), and this secondary dominant is in turn prepared by its own *ii–V* in its own key (of *V/V*, i.e. *D* major), resulting in the Em^7–A^7–Am^7–D^7 surface sequence combining chords from *D* and *G* major. Crucially, the top level *V* of this constituent is itself transformed by a tritone substitution, prepared by *ii* in its key (♭*V*, i.e. *G*♭ major). All of this derives the constituent's surface sequence Em^7–A^7–Am^7–D^7–$A\flat m^7$–$D\flat^7$. Notably, even though the chord neighbors D^7–$A\flat m^7$ in this sequence or G^7–Em^7 at the beginning (mm. 2–3) are directly adjacent, the derivation tree underpins that they are derived from different subtrees that only connect at a higher level and, hence, these chord are *not* related in terms of a direct progression. Dependencies in cases like this cannot be well captured by strictly local models.

The setup of the middle section is similar to "Take the A-Train": within an overarching *V* constituent, *V* is prepared by *IV*, by *V/V* and *ii*, whereby *IV* is stabilized by a *ii–V–I* in its own key, and *V/V* by *ii* in the dominant key. This illustrates a typical case where a dominant preparation in terms of *IV* receives higher precedence and is deeper than *V/V* and *ii*, which is reflected in the observed chord order and in preferred modulation

Figure 7: Analysis of Duke Ellington's "Satin Doll". The analysis (dropping chord repetitions and the repetition of the A-section summarized by the triangle notation) shows properties of a phrase that lacks the initial tonic. Further note that the analysis can be performed without requiring the application of prolongation except for the very top level.



centers.[87] As before, the three sections are connected by an overarching *I–V–I* group. The triangle in the tree denotes an abbreviation for an arbitrary subtree that is not relevant for the analysis displayed. In this case, the triangle abbreviates the analysis of the third phrase, which would be identical to the first phrase. Remarkably, the expressive power of

87 For a detailed discussion of the order of predominant chords and its derivation, see also Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence."

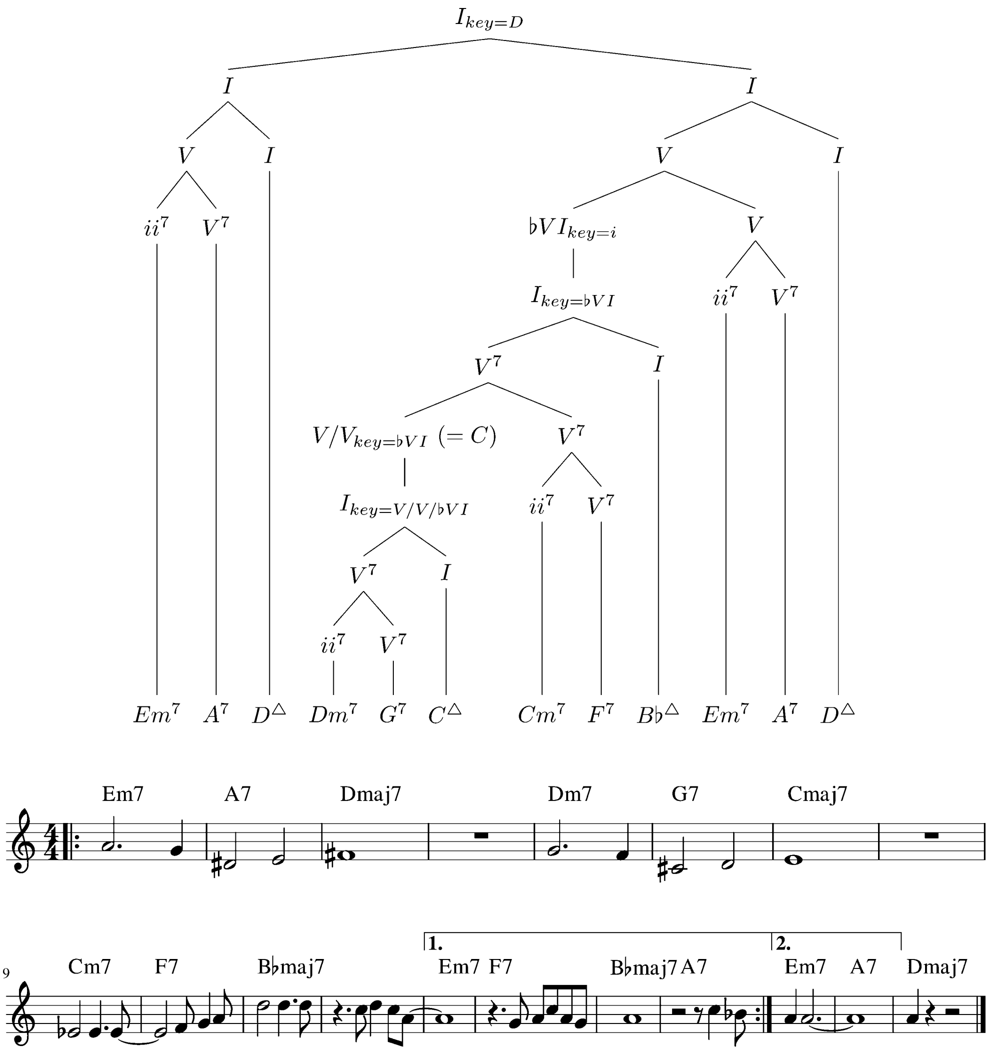
preparation, and the resulting richness in terms of tritone substitutions and tonicizations, is so strong that the derivation of "Satin Doll" does not even require the application of the prolongation rule except for the root derivation.

An interesting example for the expressive power of recursive embedding is found in Eddie Vinson's "Tune Up" (see Figure 8; the tree analysis displays only the analysis of second repetition that leads back to the tonic.). Until m. 12 the piece establishes three *ii–V–I* blocks that descend by a whole tone, tonicizing the centers $D^{\triangle}$, $C^{\triangle}$ and $B\flat^{\triangle}$ and then it reaches the dominant *ii–V* block leading back to the tonic. The blocks are not shifted arbitrarily but are all derivationally connected. The first block establishes the tonic *D* major. Further, after the first two *ii–V–I* blocks (mm. 5 and 9), the final chord is subsequently minorized and acts as the beginning of the next block. Since the third block ending on *♭VI* prepares the subsequent *V* head, the second and third block are connected up to the final tonic of the piece by recursive fifth relations and the *♭VI–V* preparation. The first *ii–V–I* block may be potentially understood as having a double function: it establishes the overarching tonic *D* and, at the same time, sets up the root of the secondary dominant of the the next station that arrives at *C* major, thus linking the three stepwise descending sequences to a long chain of relative fifths[88] (Figure 8 only displays the tonic statement analysis). That chain finally reaches *♭VI*, and then progresses into a *V*-block leading back to the tonic. As a result, this derivation constitutes a tail-recursive chain of four fifths (or, respectively, six fifths). An alternative analysis may understand the harmonic path as derived from an elaborated descending Phrygian tetrachord from *D* down to *A*, albeit in the context of *D* major, whereby each stage of the schema is tonicized as a major chord by its own diatonic *ii–V–I* preparation.

There are even more extended cases employing the remarkable power of recursive applied dominant chains and their elaborations: The final phrase of Victor Young's "Stella by Starlight" spans a direct path from $Em^{7\flat5}$ to the tonic *B♭*. That sequence begins with a tritone distance from the tonic and therefore employs tail-recursive dominant relations up to the sixth-degree applied dominant. Similarly, "Solar" (Figure 19) employs seven tail-recursive steps to traverse a path from Gm^7 back to *C* via the dominant tritone substitution *D♭*, and even more complex dependency chains may be attested to "Blues for Alice" (Figure 20).

88 Such a double function becomes possible through the fact that *ii–V–I* in *D* is surface-identical with *ii–V–I* of *V/V/V/V/♭VI*, i.e. *D* is also the fourth-order dominant of *♭VI*, which is derived here borrowing from the parallel minor. The modeling of a double function may require an operation beyond context-freeness such as internal merge such that the chord lower in the tree is internally merged with the (surface-identical) more stable chord deeper in the tree dropping its spell-out.

Figure 8: Analysis of Eddie Vinson's piece "Tune Up" (using the second repetition).



4.2. *Comparison of Frameworks*

The formalism proposed here extends the framework proposed by Martin Rohrmeier and Markus Neuwirth[89], but deviates in various significant ways from the earlier notation proposed in 2011.[90] An analysis of the first phrase of Duke Ellington's "Take the A-Train" illustrates the application of the two frameworks (Figure 9). In particular, the new way of analysis drops the phrase level (and shifts these kinds of relations to a different system: that of musical form). The analysis further drops the use of functional regions (such as the dominant region *DR*) and Riemannian function categories (such as *t*, *s* or *tp*). Thus the new way of analysis (top) simplifies the framework and merely operates at the level of Roman numerals throughout. This also reflects the understanding that larger phrases in pieces are governed by overarching hierarchical tonal relations, which captures more of the structure than having different phrases merely lined up. The bottom two analyses display the pure dependency relations between the surface symbols in the form of a tree and a dependency graph. Such pure dependency analyses, however, lack generalizations that could be drawn from abstractions like key invariance and others.

5. THE RELATION BETWEEN HARMONIC SYNTAX AND FORM

As we have seen, harmonic syntax characterizes harmonic dependencies and their interpretation. Form, by contrast, describes the regularities of phrases and their (hierarchical) organization as well as repetition structure in melodic, motivic or harmonic domains.[91] Hence, syntax and form serve different purposes, but they can interact in rich ways.

They are closely linked when it comes to the understanding of prolongation and key structures that govern whole phrases and even entire pieces, and often, analyses of both harmonic syntax and form share large parts of their substructure. The analysis of the relation between harmonic syntax and form shall be carried out at two levels: phrase and piece (leadsheet).

89 Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence."

90 Rohrmeier, "Towards a Generative Syntax of Tonal Harmony."

91 Felix Diergarten and Markus Neuwirth, *Formenlehre: Ein Lese-und Arbeitsbuch zur Instrumentalmusik des 18. und 19. Jahrhunderts* (Laaber: Laaber, 2019).

Figure 9a: Analyses of Duke Ellington's "Take the A Train". This figure illustrates the differences between the kind of analysis proposed here (top) and the one proposed in Rohrmeier (2011) (second from top).

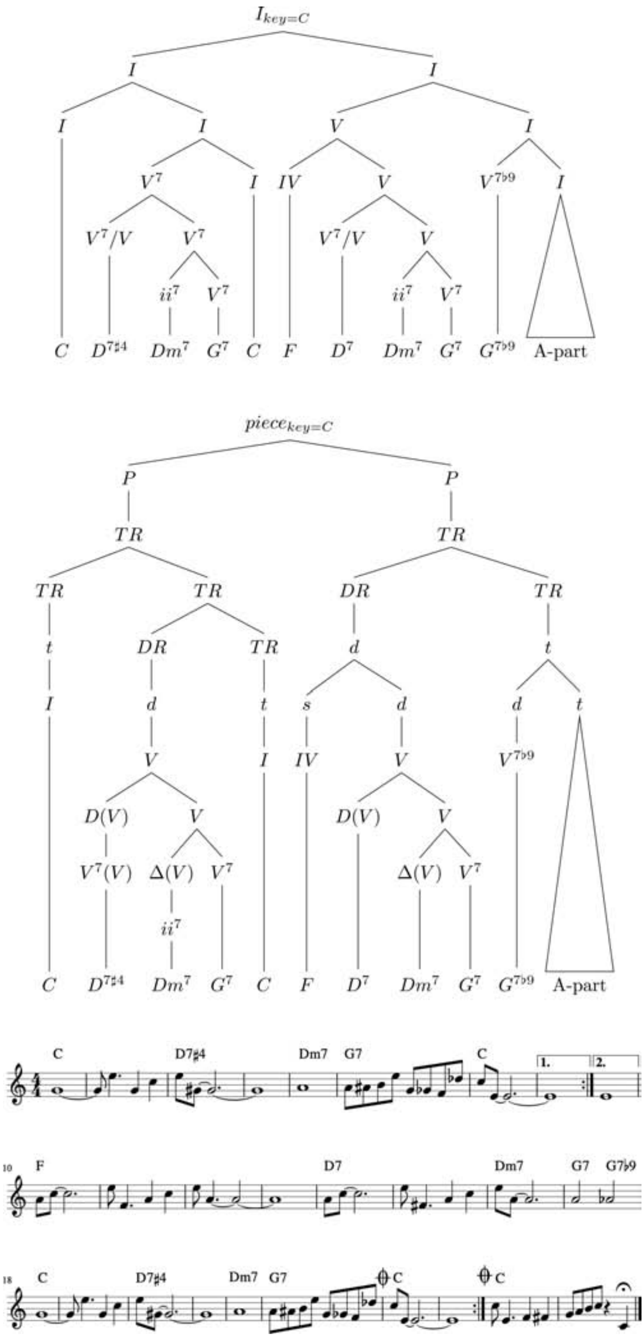
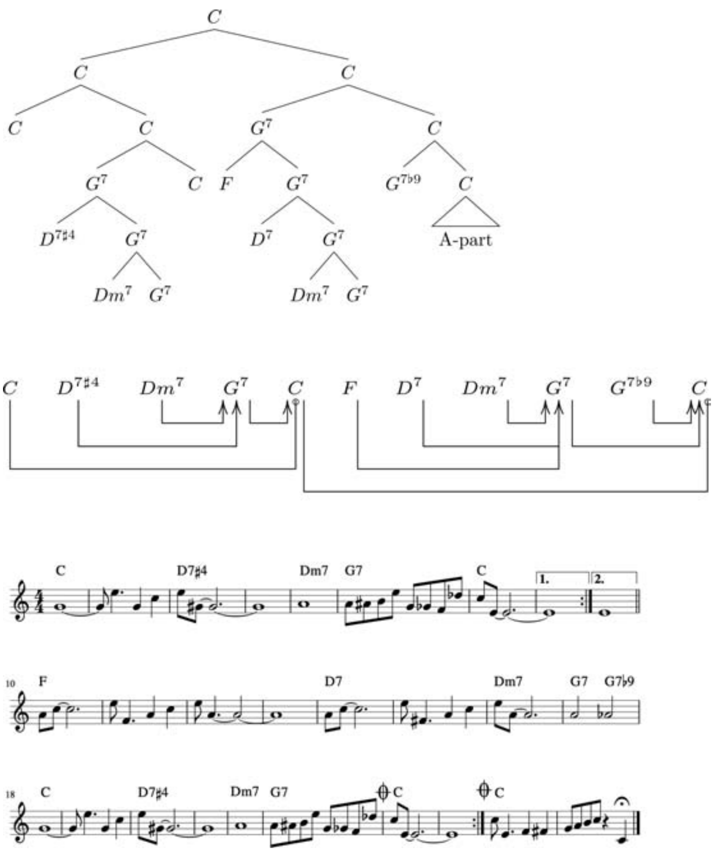


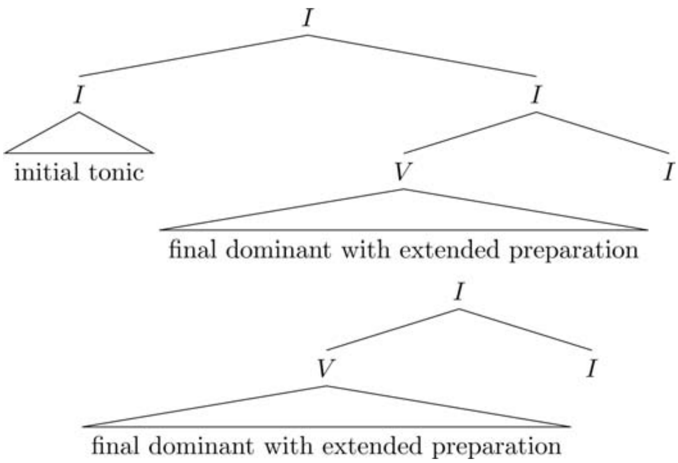
Figure 9b: Analyses of Duke Ellington's "Take the A Train". This figure illustrates a tree analysis that only notates the dependencies between the surface chord labels, not using Roman numerals and key-independent abstractions (top); and a dependency graph (above the score) that shows the harmonic dependencies between the surface chords indicating preparations (arrows) and prolongations (round ends).



5.1. *The Prototypical Structure of the Phrase*

The syntactic structure of most 8-bar Jazz phrases is generated at the top level by a simple frame of two prolonged tonic nodes, which generates the seed of the initial and the final tonic and their elaborations (Figure 10). While the initial tonic statement is commonly short, or may consist of a prolongation or dominant preparation derived from the tonic (e.g. *I–V–I* or *ii–V–I*), the path toward the final tonic through its *V* by means of recursive preparations, prolongations, and tonicizations commonly defines the harmonic body of phrases. Sometimes the recursive generation deriving from the final

Figure 10: The prototypical core phrase structure.

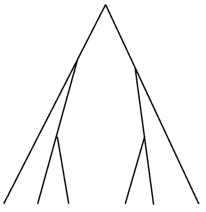


tonic is sufficiently rich that the whole phrase is generated (see, e.g., the analysis of Duke Ellington's "Satin Doll," in Figure 7, or "Autumn Leaves"[92]). Note that it is not necessary to enforce a cadence around the final tonic into the phrase template. The mere fact that the structure between the initial and the final tonic involves a recursively generated path to the final tonic through a dominant, entails a preparation by some form of primary and secondary fifths, which will, as a by-product, generate the harmonic spine of a progression that in Jazz fulfils cadential function. Among these progressions are *ii–V–I* patterns and other structures based on descending fifths and their derivations (see, for instance, the analyses in Figures 7, 18, and 20). The derivation of a high-level plagal dependency from the final tonic, instead of a dominant, lays the foundation of the Blues form (see section 5.5).

Harmony at the phrase level is often characterized as moving or jumping away from the initial tonic and "finding the way back" to the (final) tonic. Under the present formal model, this (very valid) common observation can be refined in a subtle way. This kind of "departure" (moving or jumping away) does not constitute a syntactic primitive or operation; rather, it originates from a derivation of an event from some left-recursive goal-driven preparations (which define the "path back"). The event is licensed through the final goal, not through its distance from its immediate predecessor. A derivation of jumping away is only required within the limited confinements of the derivational logic of a strictly local grammar or a Markovian model. This provides an argument why "departure" in the

92 See the analysis in Rohrmeier, "Towards a Generative Syntax of Tonal Harmony."

Figure 11: Schematic analysis of normative phrase structure as proposed by the GTTM.



sense of a Markov model or the GTTM may be dropped from a syntactic framework of music and specifically from prototypical/normative phrase structure.

Accordingly, the prototypical phrase structure proposed here suggests a revision of the "normative phrase structure" proposed by the GTTM.[93] Lerdahl and Jackendoff argue that normative phrase structure is modeled by a (prolongational) tree with balanced right branches off the top left branch (expressing departure) and left branches off the top right branch (see Figure 11 for a schematic sketch of the dependency tree). This formalization runs into problems for the reasons as just outlined: each of the right-branching chords is harmonically licensed by the chords it proceeds to rather than the chord it departs from.[94]

In contrast, the recursive right-headed derivation of the final tonic necessarily leads to the generation of chords that are remote from the initial tonic; the local contrast between the initial tonic statement and the harmonic event launching the path toward the final tonic may superficially look like a jump-away or departure, but it may rather be understood as the beginning of the subsequent left-derivation of the final tonic (see, for instance, the analyses of "Take the A-Train" (Figure 9), "Tune Up" (Figure 8), or "Solar" (Figure 19)). This observation also holds, at a higher level of abstraction, for middle sections such as bridges. Based on this reasoning, the foundation of the build-up of musical tension lies in the set of (recursively nested) goal-driven implications that are set up by such a left-derivation (rather than by the jump-away), and the release of tension corresponds to every (sub)goal that is reached in a fulfilled preparation. Such a modeling of tension naturally implies that all tension is released when the final tonic is reached. Instead of binding the concept of tension to right- and left-branching rules, the concept may be easier modeled through the interaction of the core rule types of preparation and prolongation.

93 Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*, 197–201

94 The GTTM motivates prolongational structure in terms of right-branching tension build-up and left-branching tension release. This underpinning, however, undermines the foundation that musical tension owes to the (nested) goal-driven implication-realization logic of the harmonic events involved, which requires different dependency structures.

The example of "Take the A-Train" provides a particularly challenging example for the notion of normative phrase structure as conceived in the GTTM: Lerdahl and Jackendoff posit that a full phrase would require both departure and return as specified in their normative phrase structure; particularly, a *I–ii–V–I* phrase would not suffice to constitute a complete normative phrase.[95] However, "Take the A-Train" starts off with a full phrase based on exactly such relations ($I–V^{7\sharp4}/V–ii–V–I$), in which *V* is prepared twice by *V/V* and *ii* (where an orthogonal voice leading line connects *G–G♯–A* in the middle voice). Now, if one were to consider $V^{7\sharp4}/V$ an instance of a right-branching departure of *I*, the logical link that derives the chord in the first place from the preparation of the dominant, as its label *V/V* clearly suggests, would be broken; further, a solution denying the dependency between $V^{7\sharp4}/V$ and *V* for the sake of the departure would not be consistent either: for instance, if the second chord were III^7 or V^7/vi as in ($I–V^7/vi–ii–V–I$), the same analysis as departure would apply without any trouble in the derivation even though the whole sequence would be implausible since V^7/vi would leave an open applied dominant not finding its resolution. Phrases like the beginning of "All of me" feature such an initial $I–V^7/vi$ progression, but this is only possible because it is licensed by the corresponding VI^7 chord resolving the instantiated preparation in this case. After all, a phrase built from simple preparatory relations like "Take the A-Train" highlights the need to revise GTTM's normative phrase structure and its notion of departure.[96]

Because of the tail-recursive goal-driven derivation of implicative/preparatory harmonic dependencies, it follows that the harmonic derivations need not necessarily match the symmetry that phrases commonly exhibit at the level of form. Sometimes, there is a close match, but sometimes the syntactic dependency structure reaches across (sub)phrase boundaries, thus creating a mismatch with the phrase-structure as governed by principles of form. This mismatch between syntax and form may appear relatively frequently and entails that the dependencies governing form (in terms of repetition structure and regular phrase organization) do not constitute the top-end of the tree but rather a separate kind of structure—one that needs to be coordinated with the harmonic structure in their mutually imposed constraints of harmonic anchors such as tonics, dominants, or cadences. Accordingly, the constituent structure of phrase form need not match syntactic dependency structure at the level of harmony.

One typical example of this is the first phrase of "My Way" (Figure 22). The first eight bars constitute the *presentation phrase* of the overarching *sentence* and follow a 4 + 4

95 Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*, 200–201 and figure 8.29b.

96 Also, Lerdahl and Jackendoff's example of the beginning of "La ci darem di mano" from Mozart's *Don Giovanni* exhibits a similar problem with the analysis of the ii^6 chord, which receives different readings in the antecedent and the consequent phrase, ibid., 200.

constituency grouping framed by a tonic return. The end of the first group (the *basic idea*), however, does not exhibit syntactic closure. Instead of establishing a closed constituent or a half cadence, it ends on the relative dominant of *ii* in the middle of a recursive derivation structure that reaches into the next formal group (the repetition of the basic idea). Melodic and formal boundary markers do not match harmonic syntactic structure here, illustrating that musical form and harmonic syntax are different systems.[97]

5.2. *The Missing Initial Tonic*

Classical phrases prototypically begin with an initial tonic statement or a brief cadence-like progression that states and establishes the key.[98] This is also the case for Jazz phrases. However, one may find numerous interesting cases, which dispense with the initial tonic statement.[99]

This motivates a deeper syntactic generalization: namely, that the tonic and key of a piece are not defined by the opening segment; rather, that it is the final tonic that governs the derivational logic by virtue of tonic right-headedness and finality-directed generation. As a result, phrase structures have the option to avoid the phrase-initiating tonic.

As shown before, one such case is that of Duke Ellington's "Satin Doll" (Figure 7). The example illustrates that the missing initial tonic of the phrase is easily modeled by the recursive derivation off the phrase-final tonic. In other words, instead of the overarching top-level tonic prolongation, this phrase only consists of the right-headed tonic branch, which generates all symbols on its left in the context of its preparation. Note that the initial *ii*–*V* progression of "Satin Doll" still suffices to indicate the overarching key despite the lack of an overarching tonic prolongation. A more complex case like the initial $Em^{7\flat5}$ A^7 progression in the key of $B\flat$ as it occurs at the beginning of Victor Young's

97 See also Salzer, *Structural Hearing: Tonal Coherence in Music*.

98 For instance, in the case of Mozart's Piano Sonata No. 16, K. 545 ("Sonata facile"), Gjerdingen, for lack of a syntactic category, terms this the "opening gambit" (Robert O. Gjerdingen, *Music in the Galant Style* (New York: Oxford University Press, 2007)).

99 Over the nineteenth century, the norm of the initial tonic statement is gradually loosened, for instance in compositions by Chopin or Schumann, indicating a trend of changing phrase structure. Classical precedents for this case are, for instance, the exposition in Haydn's Symphony No. 92/i (Markus Neuwirth, "Is There a 'Musical Task' in the First Movement of Haydn's 'Oxford' Symphony? Voice-Leading Schemata and Intrinsic Formal Functions," *Music Theory and Analysis* 2, no. 2 (2015): 194–203), or movements that even begin on the predominant, such as Beethoven's Piano Sonata op. 31 No. 3/i. Examples of Classical pieces lacking an initial tonic are discussed in great detail by L. Poundie Burstein, "The Off-Tonic Return in Beethoven's Piano Concerto No. 4 in G Major, Op. 58, and Other Works," *Music Analysis* 24, no. 3 (2005): 305–347 and L. Poundie Burstein, "Unraveling Schenker's Concept of the Auxiliary Cadence," *Music Theory Spectrum* 27, no. 2 (2005): 159–185. Off-tonic beginnings in the music of the Beatles are described in Naphtali Wagner, "Starting in the Middle: Auxiliary Cadences in the Beatles' Songs," *Music Analysis* 25, nos. 1–2 (2006): 155–169.

"Stella by Starlight" is remote from the diatonic of the overarching key.[100] Generally, the phenomenon of the missing tonic implies a different top-level phrase structure (no binary tonic branch and no symmetric normative phrase structure or departure as in the GTTM), yet it can be expressed without any necessary adaptations to the set of rules or the theory.

5.3. *…Or Even a Completely Missing Tonic?*

The case of the omitted tonic could even be pushed further. Why would one even need a tonic at all in a phrase or piece of tonal music? If tonality is understood in a more abstract way in terms of the syntactic principles outlined here, a top-level tonic reference may even be omitted.

This possibility arises from three aspects: (1) a piece may omit the initial tonic statement (as in "Satin Doll"), (2) it may omit the final tonic (as in the leadsheet of "Solar," Fig. 19), and (3) it may exclusively be based on the cyclic turnaround structure, e.g., ending on a final turnaround around the dominant. When these three aspects come together, a piece could be derived that follows the harmonic logic of a continuous preparation of the tonic, which however never appears. The harmonic structure of such a piece may be generated from a top-level *V* category. An hypothetical modification of "Satin Doll" illustrates this possibility (Figure 12). Alternatively, it is possible to derive such a structure by combining a dominant-tonic preparation with a unary rule that omits the tonic through an ε derivation, which creates an *empty* element ($I \rightarrow \varepsilon$).[101]

5.4. *The Harmonic Skeleton of Some Main Forms*

Moving from phrases to top-level form, one can observe that the harmonic spine of the most common forms of Jazz standards can be expressed by very similar top-level trees. Further, this analogy reveals the deeper generalization that the harmonic syntax of single tonic phrases is commonly similar or identical to that of entire pieces (leadsheets) at the top level of the tree, a result that highlights the profoundly recursive nature of harmonic structure and matches insights from Schenkerian theory and theories of form.[102]

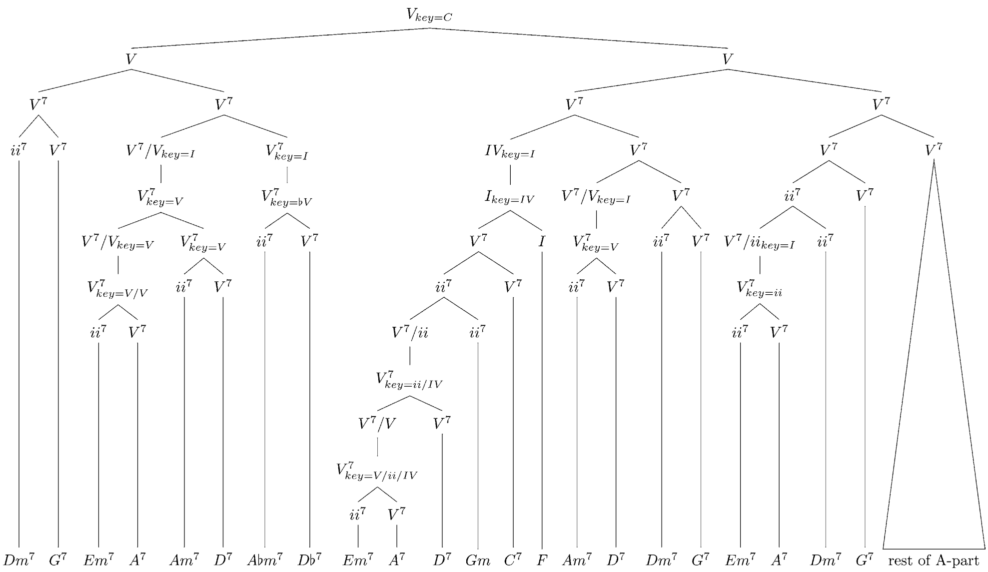
---

100 See also Terefenko, "Keith Jarrett's Art of Solo Introduction: 'Stella by Starlight'—A Case Study," for a discussion of the off-tonic beginning and an illuminating analysis of the introduction of Keith Jarrett's performance of this standard. A similarly complex off-tonic beginning can also be found with John Coltrane's "Moment's Notice."

101 However, such an empty-element rule constitutes a modification to the grammar that may result in numerous and significant unwanted side effects (see also Müller, *Grammatiktheorie*, 382–384).

102 Schenker, *Der Freie Satz (= Neue musikalische Theorien und Phantasien 3)*; Cadwallader and Gagné, *Analysis of Tonal Music: A Schenkerian Approach*; Felix Salzer, "Die Sonatenform bei Franz Schubert," *Studien zur Musikwissenschaft* 15 (1928): 86–125; Diergarten and Neuwirth, *Formenlehre: Ein Lese-und Arbeitsbuch zur Instrumentalmusik des 18. und 19. Jahrhunderts*.

Figure 12: Analysis of a hypothetical adaptation of "Satin Doll" such that there is no overarching tonic.



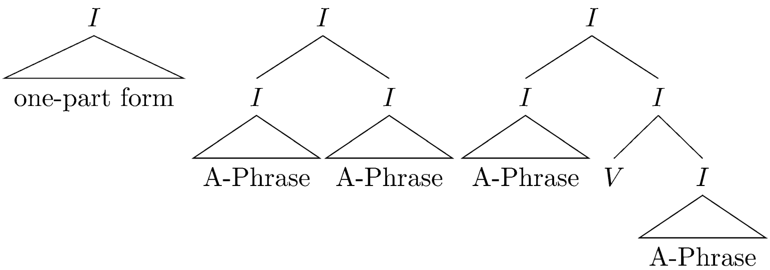
5.4.1. One-Part Form and AA Forms

One-part forms and simple AA or AA' forms most commonly exhibit very simple overarching tree templates that are fundamentally the same as those of single phrases (also, often the part exhibits only one large phrase). In the most simple cases the harmonic structure of one-part forms are entirely generated from a single overarching tonic seed. In AA forms, this top-level tonic seed is prolonged to establish the two single seeds of each A phrase (see Figure 13). An example of the analysis of a single part form is given by "Solar" (Figure 19).

One frequent point when dealing with less simple cases lies (again) in the relation between syntactic dependency and form. In an AA' schema or, more generally, forms that connect different phrases, the first phrase might embody a turnaround or lead to a *V* that establishes the second phrase (e.g. "Tune Up," Figure 8). While this *V* or turnaround occurs within the timespan of the first phrase, in terms of the dependency logic, it is derived from the tonic of the subsequent phrase, thus establishing a small conflict between grouping/form and dependency structure.[103] In fewer cases, the second phrase (A') does

103 The resulting problem of reconciling grouping, constituency (open constituents), and dependency structure will be addressed in future work. For modeling the turnaround as a cyclic tree, see Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax."

Figure 13: Syntactic frames for the core phrase structure



not begin with a tonic but a different chord (tonic preparation or missing tonic). In this case the first phrase might at its end establish a short applied dominant preparing the initial non-tonic harmony of the second phrase (harmonic up-beat). Alternatively, the first phrase may establish a final dominant preparing the head *I* of the subsequent phrase. These options may even be combined as at the end of the middle section of "Satin Doll" (Figure 7). If the turnover dominant or the "harmonic upbeat" is itself prepared extensively, a major part of the previous phrase may, in terms of the syntactic dependency logic, be associated with the subsequent phrase. All of these cases exemplify a local (or not so local) mismatch between form and syntactic structure, corroborating the conceptual difference between syntax and form.

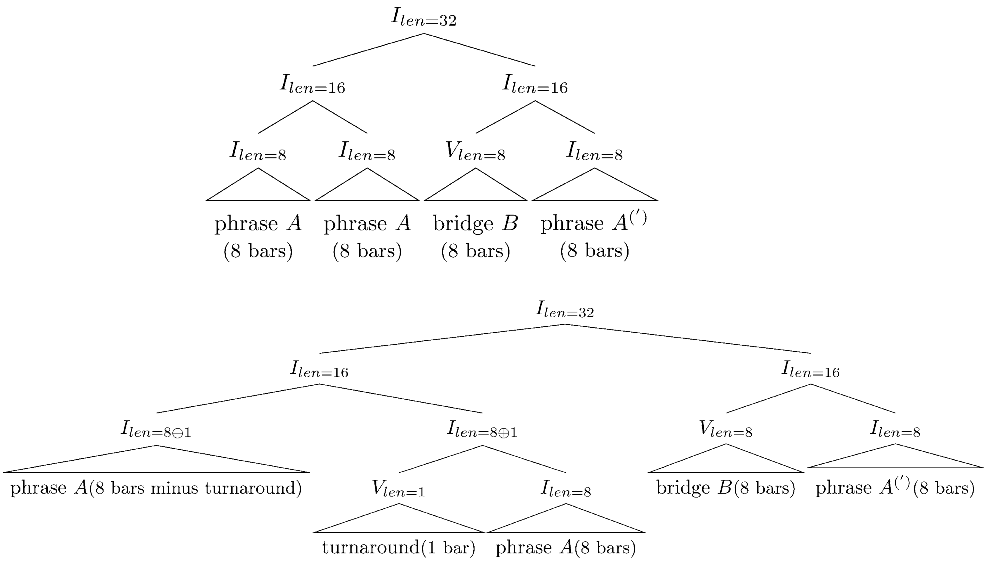
5.4.2. 32-bar AABA Form

The 32-bar AABA form is probably the most common form for classical Jazz standards, as, for instance, in terms of the foundational "Rhythm changes" schema originating from George Gershwin's song "I Got Rhythm." It commonly exhibits 8-bar A sections that frame the tonic (the first one may end on a turnaround), while the B section typically establishes a path to the dominant leading back to the final A section (for this path, often employing a descending-fifth sequence). Figure 14 illustrates a tree template for the 32-bar AABA form that expresses these observations. This schematic tree also employs the notation for harmonic rhythm (which we proposed in recent computational work[104]) in order to express the constraints concerning constituent length and subdivision of the harmonic constituents, as well as their length modifications in the case of a turnaround.

Two examples illustrating the analysis of the 32-bar AABA form can be found in "Take the A-Train" (Figure 9) and "Satin Doll" (Figure 7). A piece that maximally abstracts the

104 Harasim, O'Donnell, and Rohrmeier, "Harmonic Syntax in Time: Rhythm Improves Grammatical Models of Harmony."

Figure 14: The syntactic structure of the 32-bar AABA form as a prototype and an examples of a variant with a turnaround of one bar.



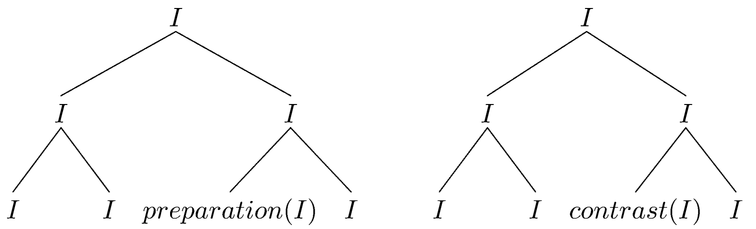
notion of large-scale AABA form is Miles Davis's "So What" (Figure 15) as well as, very similarly, John Coltrane's "Impressions." Here, the three sections are found to be just reduced to one single chord that constitutes the head. The middle section still establishes a dominant preparation; namely, it features a prolongation of the parallel minor of the tritone substitution of the dominant.[105]

For a more generalized understanding of middle sections and bridges, it might be useful to conceptualize them either as a prolongation of a preparation leading back to the head *I* of the final phrase, or as a prolongation of an overarching contrast that is established in relation to the head *I* of the final phrase. Figure 16 illustrates these two possibilities. Inherited from tradition, the middle section may use a dominant, establishing an overarching preparation of some later tonic return (Figure 16, left-handside), but it can also use different types of harmonic elaboration creating a contrast to the tonic (Figure 16, right-hand side). Such an understanding of the middle section as contrastive turns out to be useful when middle sections are comprised of hexatonic

105 In this piece, all the harmonic events are instantiated at the surface in the Dorian rather than the Aeolian mode. This is an example of the common case, in which relations between tonal centers are derived in terms of the logic of major and minor and subsequently instantiated with different modes, whereby Dorian may act like minor, and Mixolydian or Lydian like major.

Figure 15: Analysis of Miles Davis's "So What". The analysis highlights that, taking into account the octatonic / tritone substitution of *V*, the piece reflects the core skeleton of the prototypical model of the 32-bar AABA form, in which the middle section constitutes a prolongation of a dominant.

Figure 16: Two variants of the syntactic structure of a middle section or bridge.



relations (such as hexatonic poles) or stacks of fourths or fifths (the third Tonfeld), both of which will be elaborated in a subsequent article on extended tonality.

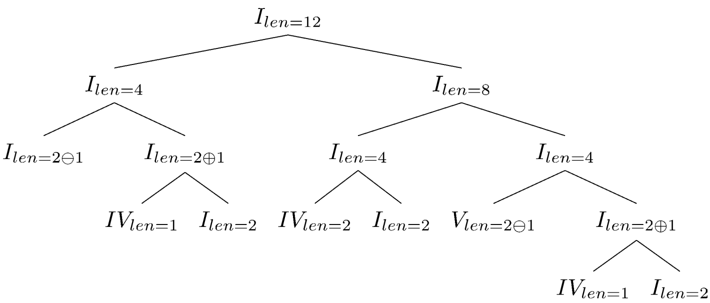
5.5. *The Blues Form and the Jazz Blues*

The Blues and the Blues form have a special status that mainly results from the central use of the subdominant *IV* in the form outline of the piece. The main features of the Blues form may be summarized as follows: the 12-bar form is divided into three groups of 4 bars. The first one establishes an initial tonic statement and a tonic prolongation, the latter often achieved through a plagal tonic-subdominant-tonic progression (with the subdominant entering in bar 2). The second 4-bar group establishes a strong subdominant in the beginning and a plagal-progression returning to the tonic (sometimes, the first phrase is regarded as a miniature instance of the first two groups). The third 4-bar group features a turnaround that first begins with the dominant, and then proceeds via a plagal progression back to the tonic (violating common-practice and Jazz preferences towards authentic harmonic progressions). It is also common that the first and second groups feature a melodic motive that is repeated in (often higher) transposition.

Figure 17 illustrates a model of the prototypical structure of the 12-bar Blues form and its syntactic analysis. In contrast to the standard harmonic language in Jazz, the Blues makes frequent and pronounced use of the plagal elaboration, and the corresponding *I → IV I* rule introduced above. The three plagal *IV–I* progressions can only be analyzed with the plagal rule *I → IV I* (maintaining a left-branching dependency relation). The *IV* becomes a central harmonic counterweight, to the extent that one might understand the deep-level *IV* as a kind of syntactic *contrast* to the tonic. Rock music shares these features and also frequent subdominant occurrences,[106] arguably because it adopted the centrality of the subdominant as a harmonic function independent of the dominant from Blues forms, as part of a general tendency towards plagal harmonic relations. Another feature

106 E.g., Temperley, *The Musical Language of Rock*; De Clercq and Temperley, "A Corpus Analysis of Rock Harmony."

Figure 17: The prototypical schema of the 12-bar Blues form and its syntactic analysis.



1:	I	IV	I	I
2:	IV	IV	I	I
3:	V	IV	I	$I\ (V)$

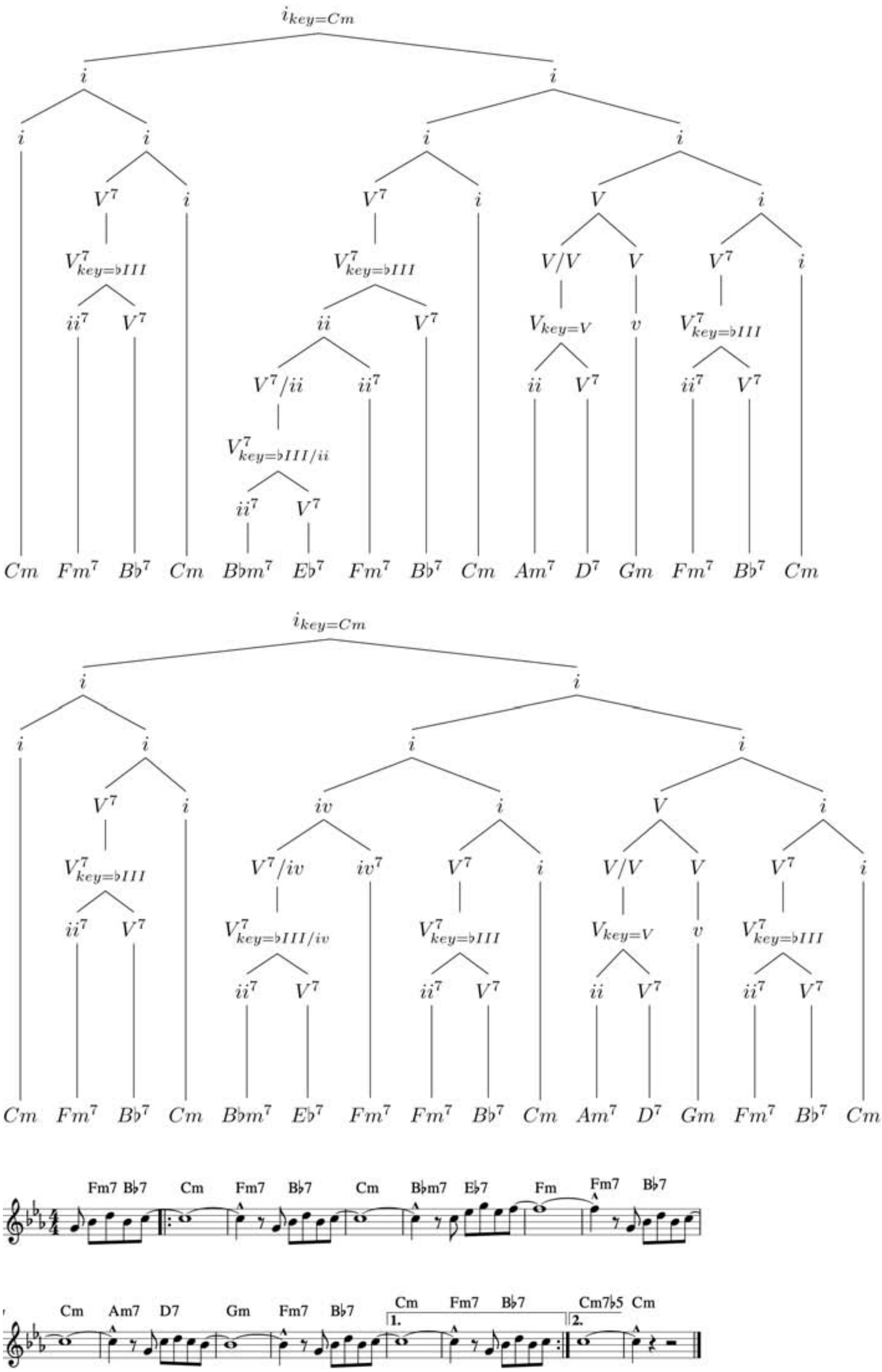
of the Blues form is that the typing constraints for the chord lexicon are relaxed, and therefore the tonic or subdominant may feature a minor seventh in the form of a I^7 chord without justifying an interpretation in terms of dominant seventh chords (or dominant functions).[107]

A piece that follows the syntactic schema of Figure 17 closely is Thelonious Monk's piece "Blue Monk." It differs, however, with regard to the penultimate *IV–I* progression, which is omitted and thus leaves only *V–I*. This fact points to a specificity of the Jazz Blues: the magnetism of implication-driven, authentic-step dominant harmony (as it is preferred in Jazz) and its resulting *ii–V–I* progressions is so strong in Jazz that it may weaken, hide or override the strong plagal progression of the traditional Blues form by variants of dominant progressions—even up to the point of almost dissolving the prototypical form. Several examples shall illustrate this point.

Miles Davis's famous "All Blues" exhibits standard Blues form; however, it employs a turnaround of V^7–$\flat VI^7$–V^7–I^7 which establishes a distinctive ambiguous hearing of $\flat VI$ between the relative of the subdominant and the tritone substitution of the secondary dominant of *V*. The final plagal turn is dropped in Monk's "Straight, No Chaser", where the last four bars feature a chain of preparations (ii/II–V^7/II–ii–V^7–I^7) instead of the typical Blues turnaround V^7–IV^7–I^7. Going even further, John Coltrane's "Blue Trane" overrides the independent plagal relations from the Blues structure (Figure 18, upper analysis). The sequence takes advantage of the fact that *iv*–$\flat VII$ constitutes a substitute of the *ii–V* preparation (the "backdoor dominant," see rule (19)). Accordingly, every occurrence of

107 For the typing constraints see also Rohrmeier, "Towards a Generative Syntax of Tonal Harmony."

Figure 18: Two analyses of John Coltrane's piece "Blue Trane".



iv can be analyzed such that *iv* is not derived from the plagal *iv–I*, but from *iv–♭VII–I* or, accordingly, $ii_{key=♭III}$–$V_{key=♭III}$–*i*, (i.e. as a predominant-dominant preparation), which makes it possible to drop the use of the plagal *I → iv I* rule in the analysis. Consequently, while *iv* still appears on the surface at mm. 2, 5, 6, and 10 at the same positions as in the standard Blues schema, and *iv* directly follows *v* in mm. 9–10, the analysis reveals that the logic of the underlying deep-structure derivation employs only dominant preparations, especially in mm. 9–10, where the *v–iv* adjacency is 'suspended' at the higher level[108]. Notably, the middle four bars could also be analyzed using the more traditional schema and having *iv* as a second stable syntactic point (bottom analysis in Figure 18). However, it is crucial to bear in mind that syntactic derivations and musical form are different systems; so a "hearing" of *IV* as the strong blues subdominant at the level of form need not necessarily be reflected at the syntactic dependency level. These differences will become even more important in light of the subsequent more complex examples.

5.5.1. Is "Solar" a Blues? Is "Blues for Alice" a Blues?

How far can one go with the modification of the Blues schema? In an oft-cited article, François Pachet examines the question whether "Solar" is a Blues from a computational/syntactic modeling perspective.[109] Because of some harmonic surface similarities to the Blues schema, this piece provides an interesting case (even if the difference from the Blues strikes immediately); and it serves as an example raising questions concerning the criteria that define the Blues form and their implications for the general understanding of musical form.

At the surface level, "Solar" indeed exhibits some similarities to the Blues schema: it is 12 bars long, and features a tonic prolongation in the first four bars followed by a strong tonicized subdominant at the beginning of m. 5. The third 4-bar group is a prolongation of the dominant in terms of an extended turnaround the preparation of which, in fact, started much earlier. However, as a detailed analysis (Figure 19) reveals, the subdominant is derived not from a plagal subdominant-tonic dependency (especially given the missing final tonic), but from a sequence of deep recursive fifth-relations that govern the entire structure of the piece along with a slowly descending melodic anchor line that descends every two bars and accelerates at the end (*C* down to *F*). Also the third 4-bar group establishes neither a final tonic nor a clear dominant constituent at mm. 9–10, but rather

108 This analysis highlights that a local *V–iv* or *V–IV* progression may not at all be ungrammatical or underivable with the logic of the dominant-driven diatonic grammar. The plausibility of the sequence is defined by the subsequent chords, as, for example, in *V–iv–♭VII–I* or *V–IV–♭VII–I*.

109 Pachet, "Computer Analysis of Jazz Chord Sequences: Is Solar a Blues?"

Figure 19: Analysis of Miles Davis's "Solar" (the final tonic is left empty (ε) in order to express the cyclic structure of the turnaround).

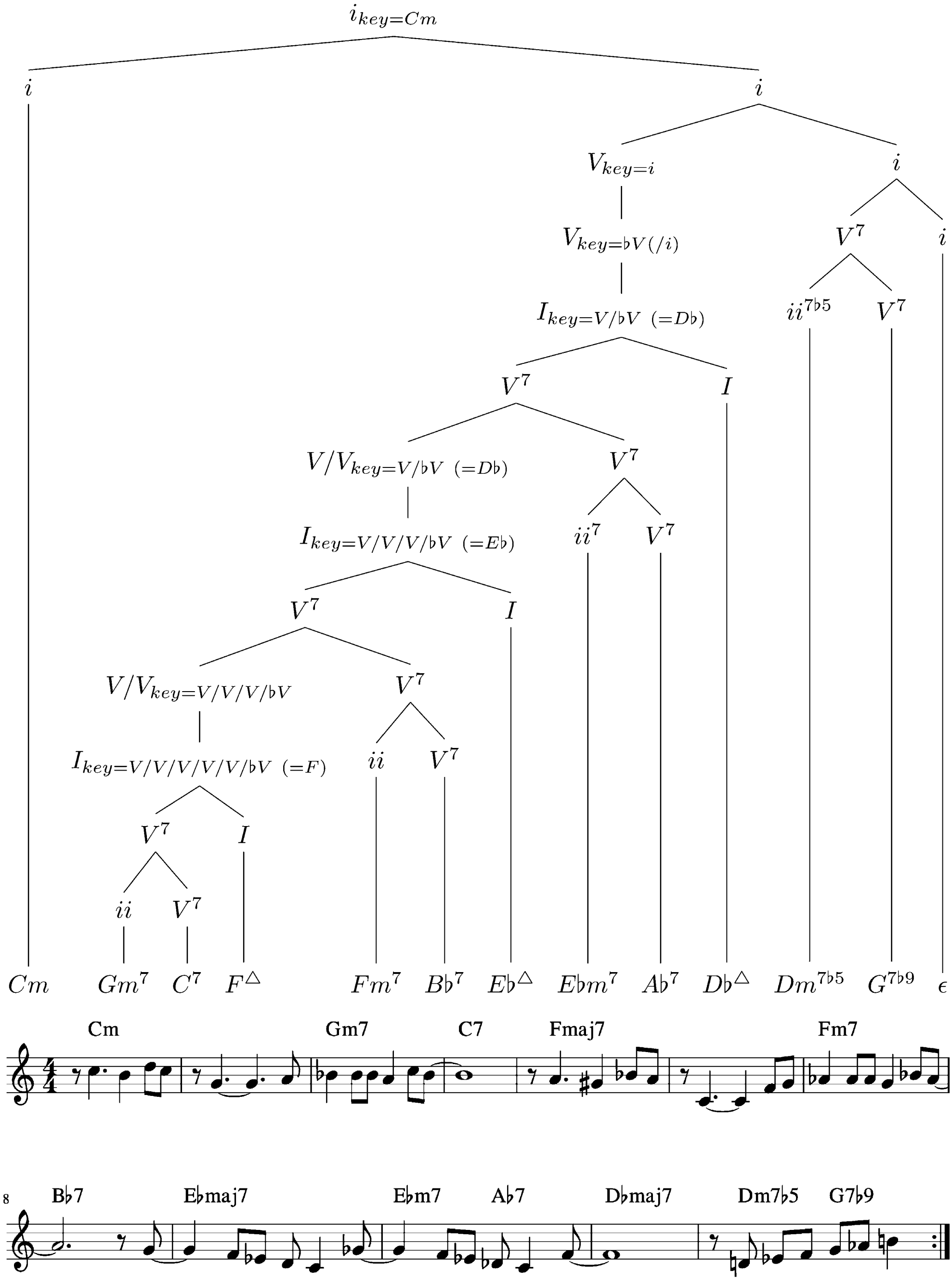


Figure 20a: Two analyses of Charlie Parker's "Blues for Alice".

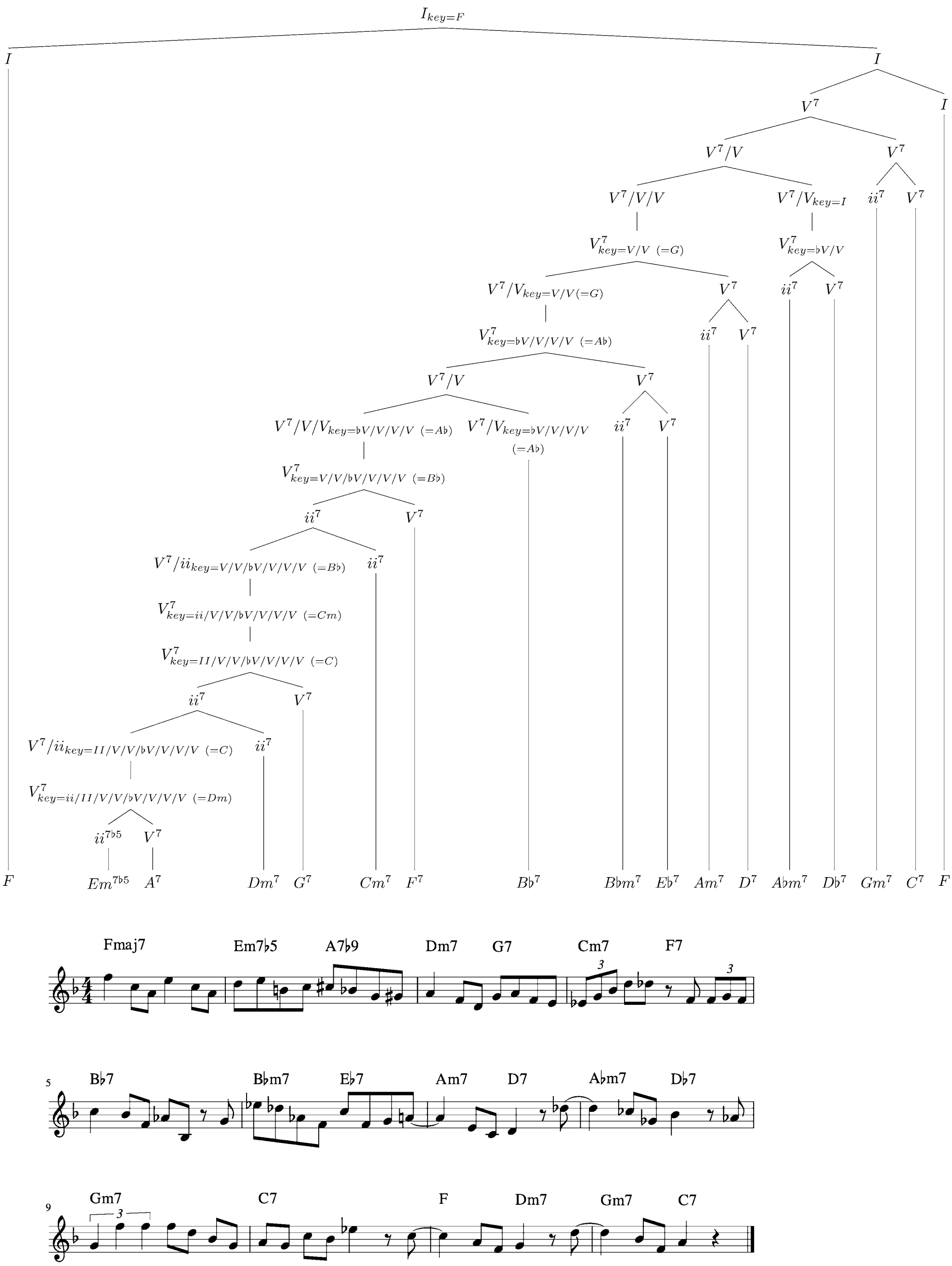
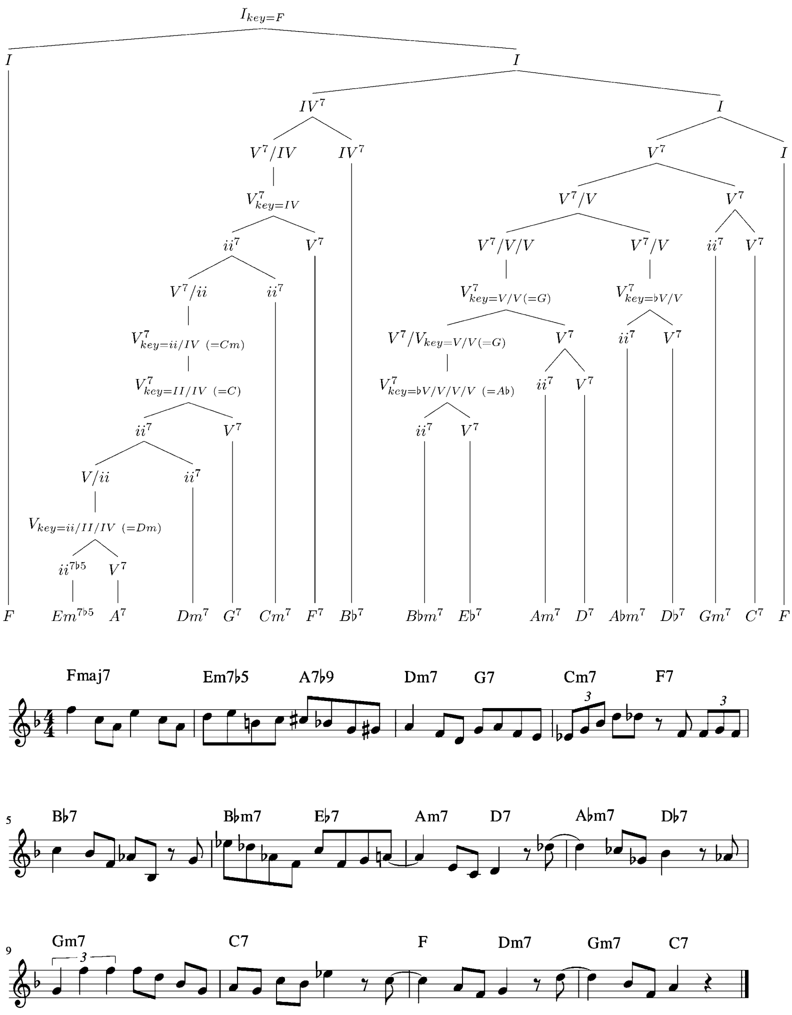


Figure 20b: Two analyses of Charlie Parker's "Blues for Alice".



the *E*♭△ chord (which may, if anything, instead induce relative tonic associations). Hence, the derivational analysis of "Solar" reveals a dependency structure that is significantly different from that of the Blues schema. Despite some surface similarities (in particular a strong *IV* in m. 5), the deep structure analysis justifies the rejection of the assumption that "Solar" is derived from the same underlying Blues schema as outlined above and in Figure 17, since it lacks, for instance, the tonic return in mm. 7–8 or mm. 11–12, and because the descending melodic line does not evoke schematic Blues lines and scale degrees either.

Given this analysis of "Solar," another case is also thought-provoking, namely, Charlie Parker's Bebop piece "Blues for Alice" (see Figure 20a and b for its analysis). Like "Solar" and "Tune Up" (Figure 8), the main harmonic body of the piece is derived from a left-recursive derivation of multiple applied dominant relationships (and tritone substitutions thereof) with chords drawn from their respective keys. The outline of the piece follows the Blues schema such that the subdominant *B♭* is established in m. 5, the "tonic degree" is touched upon in passing (in the form of an applied dominant) in m. 4, and a turnaround to *F* is established in m. 9 as *ii–V–I*, reaching the tonic in m. 11. The first four measures immediately leave the initial tonic *F* in order to prepare the $B\flat^7$ chord in m. 5. This is achieved through a chain of three applied *ii–V* blocks, whereby one block prepares *IV* and the other two each recursively prepare the *ii* of the subsequent one. Altogether, this results in a sixth-degree fifth distance traversed from the (out-of-key) chord $Em^{7\flat5}$ up to *B♭*. The analysis of the $B\flat^7$ chord in m. 5 is, in turn, ambiguous. While it may be derived as a subdominant branching off the top-level *I* or the top-level *V* (see the analysis in Figure 20b), it also links with another long chain of fifths leading back to the dominant C^7 (m. 10) and, ultimately, the tonic *F* (m. 11; see the analysis in Figure 20a). More specifically, it functions as an applied dominant to $E\flat^7$, which is the applied tritone-substituted dominant to D^7, which in turn is the applied dominant to Gm^7 at the beginning of the *ii–V–I* return to *F*. This applied dominant connection is interspersed with another applied *ii–V* of $A\flat m^7$–$D\flat^7$ between D^7 and Gm^7. Accordingly, the harmonic sequence establishes a link between $B\flat^7$ and Gm^7 in terms of a triple applied dominant relation. Since $B\flat^7$ is in turn linked to the second chord $Em^{7\flat5}$ (m. 2), the entire sequence from the second chord up to the final tonic can be analyzed as linked by a chain of recursively applied dominants up to the eleventh degree. The derivation takes advantage of the facts that I^7 is surface-equal with V^7/IV and that the subdominant IV^7 is surface-equal with the secondary dominant of the tritone substitute of the third-order dominant (i.e., a far reaching application of tail recursion, expressed as $V^7/V/\flat V/V/V/V$ in the upper figure); the piece spaces the chords in such a way that they appear on the locations predicated by the classical Blues schema. Similarly, the derivation of "Solar" takes advantage of the fact that *IV* is surface-identical with *V/V/V/V/V/♭V*. Note that both derivations involve six derivation steps including one tritone substitution for the 'dominant algebra' to work out, yet the substitution appears at a different point in the derivation. Generally, these two pieces demonstrate once more the enormous generative/expressive power of relative dominant relations and tritone substitutions to derive complex sequences and to draw relations between remote keys, which is a typical feature of advanced Jazz harmony.

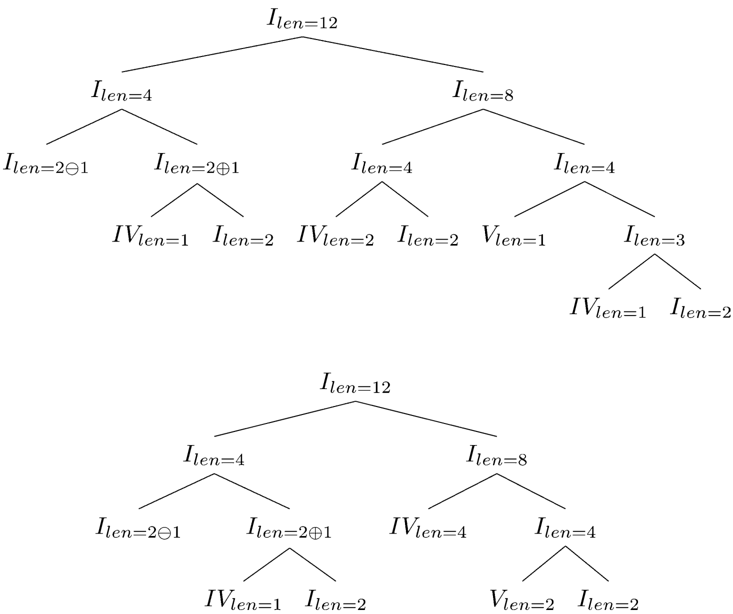
After all, where does this leave us with the Blues schema? Let's look back at the examples discussed and summarize. Looking at the turnaround, "Blue Monk" and "Straight, No Chaser" follow all features of the Blues schema except for the third 4-bar group featuring the *V–IV* turnaround. "All Blues" substitutes the subdominant for a *♭VI* in the turnaround.

"Blue Trane" features a *v–iv* progression which is, however, dissolved by overarching authentic progressions (*v–i* and *iv–♭VII–i*) and may rather be regarded as an extension of the final tonic constituent. Also, the other examples feature no *V–IV* turnaround group. Regarding the initial 4-bar tonic or *I–IV–I–I* group, all examples exhibit this feature—taking into account that the preparation of the second subdominant 4-bar group may reach back into the first group. In "Blues for Alice" the latter leads to a maximal abstraction of the first group being reduced to the initial tonic only briefly stated in m. 1 (and, as its surface form, briefly passed as *V/IV* in the second half of m.4). According to the schema, the second 4-bar group establishes a strong subdominant at its beginning, returning to the tonic with a plagal step before the turnaround of the third group. "Straight, No Chaser" and "All Blues" conform to this, while "Blue Monk" inserts *IV–I°⁷–I* into the plagal step and "Blue Trane" transforms the plagal step into the authentic *iv–♭VII–I*. "Blues for Alice" skips the tonic return and proceeds from *IV* via recursive authentic steps to *V*.

These variations of the second 4-bar group entail that there is no single, fixed abstract Blues schema in terms of a top-level dependency tree the leaves of which generate all the versions: versions with a plagal tonic return and versions that have no tonic in the second group cannot be derived from the same underlying tree structure. Accordingly, the Blues form may rather be required to be tied to certain harmonic-metrical features: an initial 4-bar tonic constituent, often with *IV* in m. 2, a second 4-bar group that features a strong *IV* on its beginning, and a third 4-bar group that features an elaborated *V–I* group, with the head *V* of the dominant constituent being established in mm. 9–10 and the tonic constituent in mm. 11–12. This appears to be the smallest common denominator of the examples discussed. Another general observation is that while traditional Blues forms may be convincingly—and necessarily, when additional bridging chords are missing—derived from plagal derivations of the tonic, the *Jazz Blues* often follows a different logic whereby, as is typical of Jazz, the plagal progressions are avoided at the surface and deep levels, being replaced by variants of *V–I* and *ii–V–I* progressions that bind the subdominant chords to overarching dominant progressions, as in the example of John Coltrane's "Blue Trane." This concerns the turnaround in particular, where the *V–IV–I* progression is commonly replaced by some variant of a *ii–V–I*.

From all of the variants discussed above, one might draw two different conclusions regarding the form of the Jazz Blues. According to one conclusion, facing the large variance of different (Jazz) Blues pieces, one may need to resort to a characterization of the form based on a certain bundle of loosely shared harmonic-metrical surface features. Alternatively, one may stipulate different variants of the overarching tree schema for the (Jazz) Blues form. In addition to the basic traditional Blues schema and its analysis (Figure 21, top), for instance, a second, more specialized schema for the Jazz Blues (Figure 21,

Figure 21: A syntactic characterization of the 12-bar Blues schema by two prototypical tree templates.



bottom) needs to be assumed to account for the examples discussed here. The assumption of these schemata allows one to draw the distinctions that explain the range of deviations of a piece like "Solar" that make it immediately sound very different from the Jazz Blues schema, whereas "Blues for Alice" still fulfills the above criteria, despite all of its complexity.

Another strategy to model the Blues form consists of defining the Blues as a fixed top-level schema with a flat derivation of the *IV*, the *V*, and the turnaround, which is then subject to subsequent, more surface-level derivations and ornamentations based on rewrite operations. Mark Steedman and Jonah Katz each propose an analysis in this vein (rule 0 in Steedman's proposal).[110] Steedman further includes a right-branching rule $X \rightarrow X\ IV/X$, which stands in contrast to the set of rules proposed in this article. Katz also includes such a rule in his grammar and derives *IV*–♭*VII* (as in "Blue Trane," Figure 18) as two recursively-nested forward-driven subdominants (subdominant reaching out to the subdominant of the subdominant).[111] The grammar presented in this article avoids such forward-propagating (plagal) analyses (and, in general, right-branching rule types) because

110 Steedman, "The Blues and the Abstract Truth: Music and Mental Models"; Katz, "Harmonic Syntax of the Twelve-Bar Blues Form: A Corpus Study."

111 Katz, "Harmonic Syntax of the Twelve-Bar Blues Form: A Corpus Study."

of its lack of interpretation and its mismatch with the goal-driven logic of functional harmony. In contrast to Steedman and Katz, this article proposes to understand ♭*VII* as an authentic substituted ("backdoor") dominant preparation of *I* (as in Figure 18), and occurrences of *IV–I* as left-branching plagal dependencies.

6. GENERAL DISCUSSION

The assumptions above, the core rule types, and the rules to model out-of-key chords and modulations are argued to characterize the foundation of diatonic syntax. The specific rules proposed here advance previous work by other scholars and myself,[112] yet with differences in the specifics. These contributions involve the proposal of the core rule types, the foundation in the diatonic system, the way to model out-of-key derivations and modulation with its recursive tonic embedding, as well as the specification of the relation between overarching tree structures and phrase and form templates.[113]

6.1. *Locating the Theory*

An essential aspect of the proposed syntax theory—in contrast to many linguistic theories of syntax—is that it reflects the fundamentally *temporal* nature of music. The core rule type of preparation characterizes the goal-driven implication-realization logic of harmonic dependencies that is tied to temporal structure. While the order of words or constituents may be flexible and subject to variation within or between different languages, the implication-realization logic is fundamentally left-branching and right-headed.[114] Similarly, prolongation is an inherently temporal concept that captures the extension of a harmonic object over a timespan. It may be owing to this temporal foundation that

---

112 Steedman, "A Generative Grammar for Jazz Chord Sequences"; Steedman, "The Blues and the Abstract Truth: Music and Mental Models"; Lerdahl and Jackendoff, *A Generative Theory of Tonal Music*; Lerdahl, *Tonal Pitch Space*; Rohrmeier, "A Generative Grammar Approach to Diatonic Harmonic Structure"; Rohrmeier, "Towards a Generative Syntax of Tonal Harmony"; W. Bas De Haas, "Music Information Retrieval Based on Tonal Harmony" (PhD diss., Utrecht University, 2012); Tojo, Oka, and Nishida, "Analysis of Chord Progression by HPSG"; Granroth-Wilding and Steedman, "Statistical Parsing for Harmonic Analysis of Jazz Chord Sequences"; Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax"; Katz, "Harmonic Syntax of the Twelve-Bar Blues Form: A Corpus Study."

113 While many of the syntactic rules are similar to the ones by Granroth-Wilding and Steedman (as they note as well in their paper), differences lie, for instance, in the assumption of backward dependencies (such as $I \rightarrow I\ IV$) or in the way of modeling modulation. The proposed syntactic analysis of the overarching form derived from a single tonic constitutes a main difference from the grammar proposed by Granroth-Wilding and Steedman, who employ a development rule that combines any complete constituent – a rule which is overly unconstrained (Granroth-Wilding and Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences").

114 This way, this formalization is related to approaches to music analysis based on modal temporal logic, such as Alan Marsden, *Representing Musical Time: A Temporal-Logic Approach* (Swets & Zeitlinger Publishers, 2000).

harmonic syntax is closely correlated with metrical structure and harmonic rhythm[115]—a result corroborated by computational modeling.[116] The formal theoretical underpinning of the interplay between these structural domains will be further examined in future work on generative musical rhythm.

The essence of the proposed syntactic theory, specifically the assumption of the core syntactic rule types—preparation and prolongation, and substitution—as well as a precise theorizing of diatonic harmony, modulation, and borrowing, all constitute major differences to the GTTM. As argued above, in the presented theory, *departure* is unnecessary for modeling syntactic relations because it lacks a corresponding interpretation, and can be dropped from the syntactic framework (not only for Jazz but also for Classical music). As a consequence, the prototypical phrase structure as formulated in the GTTM requires revision, as argued in the section on phrase structure (5.1). The importance of the plagal rule (section) further suggests that a dominant cannot be identified through its left-branching derivation from a stable tonic. The grammar requires the assumption of categories and two separate rules distinguishing authentic preparations from plagal relaxations.

The formalization of harmonic syntax with the outlined principles is not specifically bound to abstract context-free grammars and it could, for instance, be formulated in terms of lexicalized grammars. An adaptation in terms of a minimalist framework[117] would employ *external merge* over instances from the chord lexicon. The formalism could be also expressed with *categorical combinatory grammar* (CCG) as advanced by Mark Steedman[118], where instead of generative rules possible chord functions would be modeled by a rich lexicon and be combined with combinators such as function application, composition, coordination and development.[119] Challenges with such (lexicalized) adaptions may lie in the formalization of unary rules, key features and its casting as they occur specifically with substitution, out-of-key chords and modulation. While parsing is efficiently possible with CYK parsing[120], one main challenge for an adaptation with CCG lies in modelling

115 Stefan Love, "An Approach to Phrase Rhythm in Jazz," *Journal of Jazz Studies* 8, no. 1 (2012): 4–32; Keith Salley and Daniel Shanahan, "Phrase Rhythm in Standard Jazz Repertoire: A Taxonomy and Corpus Study," *Journal of Jazz Studies* 11, no. 1 (2016): 1–39.

116 Harasim, O'Donnell, and Rohrmeier, "Harmonic Syntax in Time: Rhythm Improves Grammatical Models of Harmony."

117 Chomsky, *The Minimalist Program*.

118 Mark Steedman, *The Syntactic Process* (Cambridge, MA: MIT Press, 2000); Steedman, "The Blues and the Abstract Truth: Music and Mental Models."

119 Raymond M. Smullyan, *To Mock a Mockingbird and Other Logic Puzzles Including an Amazing Adventure in Combinatory Logic* (New York: Oxford University Press, 2000); Granroth-Wilding and Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences."

120 Steedman, *The Syntactic Process*, ch. 9.

and computing forward generation, which also may make (statistical) sampling more difficult for probabilistic implementations of the grammar and grammar inference.

Another topic that requires brief discussion concerns schema theories and analyses by patterns. Many harmonic progressions in Jazz employ Baroque and classical prototypes as known in schema theories,[121] such as the *Lamento*, *Monte*, *Fonte*, *descending fifths sequence*, and others. In a similar vein, John Elliott and Conrad Cork gathered a large number of prototypical Jazz chord patterns.[122] The identification of common patterns and their variants is a very useful resource for stylistic pedagogy, or comparative historical research, and musical schema theory has been linked to the branches of construction grammar.[123] The purpose of a syntactic theory as outlined here stands somewhat orthogonal to schema theory: The focus lies on understanding hierarchical dependencies and constituents built from chords up to the entire leadsheet at the granularity level of the single harmonic event. While schemata are mid-level building blocks, they do possess internal substructure and are also linked to the overarching context in terms of the dependencies as outlined here. Harmonic schemata may hence be regarded as fixed subtrees (treelets) that may be used as a whole chunk inside a piece or may potentially be internally modified or elaborated following syntactic rules. In other words, harmonic schemata may be conceptualized in terms of fragment grammars, for instance, as proposed by Tim O'Donnell.[124]

Similarly, even though this article focuses on modeling harmonic dependencies it is not argued that harmony is fully separable from voice leading in the way it governs musical structure.[125] Harmonic events are governed by a double determination: Their dependencies may be analyzed in terms of syntactic chord relations as proposed here, while voice leading constitutes an orthogonal dimension that regulates the linear connection between events (see, for instance, in the middle voice in "Take the A-Train", the descending line in "Solar" or the analysis of "My Way"). The integration of both in terms of an overarching syntactic framework will be tackled in future work.[126]

---

121 E.g., Gjerdingen, *Music in the Galant Style*.

122 John Elliott, *Insights in Jazz* (London: Jazzwise Publications, 2009); Conrad Cork, *Harmony with LEGO Bricks: A New Approach to the Use of Harmony in Jazz Improvisation* (Tadley Ewing, 1990).

123 Robert Gjerdingen and Janet Bourne, "Schema Theory as a Construction Grammar," *Music Theory Online* 21, no. 2 (2015); Adele Goldberg, *Constructions: A Construction Grammar Approach to Argument Structure* (Chicago: University of Chicago Press, 1995).

124 Timothy J. O'Donnell, *Productivity and Reuse in Language: A Theory of Linguistic Computation and Storage* (Cambridge, MA: MIT Press, 2015).

125 See, for instance, Aldwell and Schachter, *Harmony and Voice Leading*; Gauldin, *Harmonic Practice in Tonal Music*. For a cognitive study disentangling the effects of voice leading and harmony, see also Leona Wall et al., "The Impact of Voice Leading and Harmony on Musical Expectancy," *Scientific Reports* 10, no. 5933 (2020), doi:https://doi.org/10.1038/s41598-020-61645-4.

126 Note that because the intersection of a regular grammar with a context-free grammar is context-free, the resulting outcome of a joint theory of (local or regular) voice-leading and hierarchical harmony is again to be expected to be of (at least) context-free complexity.

The main contribution of this article is the proposal of a grammar formalism and its core principles. There is no claim for the grammar fragment of the set of concrete rules proposed here to be final or all-encompassing. For the analysis of larger or style-specific corpora, additional rules, restrictions or special cases may need to be assumed or adaptations may be necessary; nonetheless, I argue that the overarching framework encompasses the core generalizations that need to be assumed to analyze harmonic syntax.

Importantly, the analyses in this article do not claim to be the only possible hearings of the pieces, and the grammar potentially expresses a large number of different parse trees for any given piece. The ambiguity of grammatical analyses with regard to certain phrases allows for different hearings and interpretations, or superpositions of analyses to be expressed (e.g. the union of several or all analyses). The grammar rather equips theorists with a tool to notate precise interpretations and their corresponding dependency structures, such that different analyses could be compared and weighed against each other. Different parse trees may indeed reveal different tendencies of chords in the overarching forest of harmonic relations. The proposed formalism provides a formally precise notation language for different analyses, similar to other music theoretical notation languages such as Roman numeral notation, Schenkerian reduction, or a GTTM tree, but more precise and rigorously linked to mathematical and computational modeling. The grammar further comes with a natural way to understand elaborations, variation and, generally, similarity between chord sequences and their derivations. Since the proposed rules express rewrite steps that elaborate simple sequences to more complex ones, they can express precise derivations between core sequences and their derivates (e.g., different variants of *ii–V–I* sequences or different versions of a leadsheet). The grammatical parse trees of a previous form of this grammar have been employed in a computational study to model similarity between Jazz leadsheets.[127]

6.2. *The Expressive Power of the Formalism and the Grammatical Complexity of (Jazz) Harmony*

One core discussion in the community concerns the formal complexity required to be assumed for music.[128] A key question for harmonic syntax is whether harmony is strictly local, context-free, or (mildly) context-sensitive as is the case for language[129].

127 De Haas et al., "Modeling Harmonic Similarity Using a Generative Grammar of Tonal Harmony."

128 Katz and Pesetsky, "The Identity Thesis for Language and Music"; Rohrmeier et al., "Principles of Structure Building in Music, Language and Animal Song"; Jackendoff and Lerdahl, "The Capacity for Music: What is It, and What's Special About It?"; Dmitri Tymoczko, *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice* (New York: Oxford University Press, 2010).

129 Stuart M. Shieber, "Evidence Against the Context-Freeness of Natural Language," in *Philosophy, Language, and Artificial Intelligence* (Dordrecht: Springer, 1985), 79–89; Steedman, *The Syntactic Process*.

There are plenty of musical examples in this article where non-local harmonic dependencies and center-embedding are required to model the dependency structure between all chords. The analysis of "Satin Doll" (Figure 7), for instance, contains several chord adjacencies that are only explained at an overarching hierarchical level. Generally, modulations may generate center-embedding that cannot be captured with plain strict locality either. One core aspect of the generalizing power of the proposed formalism is that relations between tonicizations and larger key regions can be modeled within the same overarching generative framework without requiring to assume additional parts of the theory that handle modulations, key relations non-local phenomena or phenomena outside tonal phrases. All of these cases constitute challenges for simple strictly local or regular grammars.

The analyses presented in this article do not employ derivations beyond context-free complexity. Hence, the dependency structures discussed here do not draw on syntactic mechanisms of mild context-sensitivity. The integration of cases that involve potentially higher complexity such as the half cadence and its corresponding conflict between dependency and constituency structure[130] will be addressed in subsequent work. A discussion whether or not music, and particularly harmony, incorporates mildly context-sensitive dependencies requires an analysis of Katz and Pesetzky's argument as outlined in the following sections.

The standard "My Way" is a case well-suited for the exploration and analysis of extended plagal progressions and the use of *IV*, as well as the relation between harmony and voice-leading. In particular, the climax of the first phrase provides an interesting example (see Figure 22). The second part of the phrase exhibits an extended plagal return to *I*, which is itself productive and becomes prepared by its relative *ii*–*V*. Furthermore, the plagal progression is derived twice, first from the original major diatonic and second as borrowed from the minor diatonic, while there is a linear chromatic descent in the middle voice. A similar case is found by *ii* and $ii^{7\flat5}$ at the beginning, both being derived as preparations of the dominant of the superordinate *ii*. Both chords each exhibit the harmonic dependency as preparation of the head, while, at the same time, they exhibit a linear chromatic voice-leading transition. It is not useful to model the chromatic alteration as a dependency of the chord of origin (here, of *ii* or *IV*) because this would loosen the relation that the head of each is *V* and that either could be dropped (deletion test). While the voice-leading connection is local, the harmonic dependency relation is hierarchical in regards to the goal chord. This provides an example for the double determination of chords from the complementary systems of harmony

130 Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence."

and efficient voice leading. Finally, the interpretation of the derivation of the second chord is not a *iii* departing from *I* (no departure rule necessary), but a Dorian *ii* of the dominant *V* (in the key of *ii* Dorian, which is the mode of *ii* that matches the scale of *I* and has therefore its *ii* match with *iii* in the key of *I*). Another case is manipulation of the *ii* chord through the chromatic 8–♯7–♮7 line in the middle-voice over the *Dm* prolongation (mm. 5–7), which is not analyzed in Figure 22.

The strong subdominant established in the second part of this phrase, as well as the plagal dependencies in Blues forms underpin the requirement of an independent plagal progression rule $X \rightarrow IV/X\ X$, even though this rule may occur rarely in Jazz as pointed out earlier. Further, plagal progressions can even be inserted between dominant and tonic (see the end of the phrase), reminiscent of the previous strongly pronounced subdominant. This interrupts the strong dominant-tonic bond and introduces a final plagal turn just before the final tonic. The subdominant insertion provides potential counterevidence concerning Katz and Pesetsky's internal merge argument, which stipulates a "lock" between the final dominant and the tonic (such that no other functional derivation can enter in between).[131] Other examples of similar cases can be found in the plagal and double plagal progressions in Rock music,[132] such as "Let it be" by the Beatles. Katz and Pesetsky note that they are aware of such issues, arguing that they would, in the terminology of this article, constitute "improper harmony" that lies outside the syntactic system. Therefore, the debate rests on whether to understand such plagal insertions and progressions as instances of functional harmony as opposed to mere voice-leading appogiaturas.

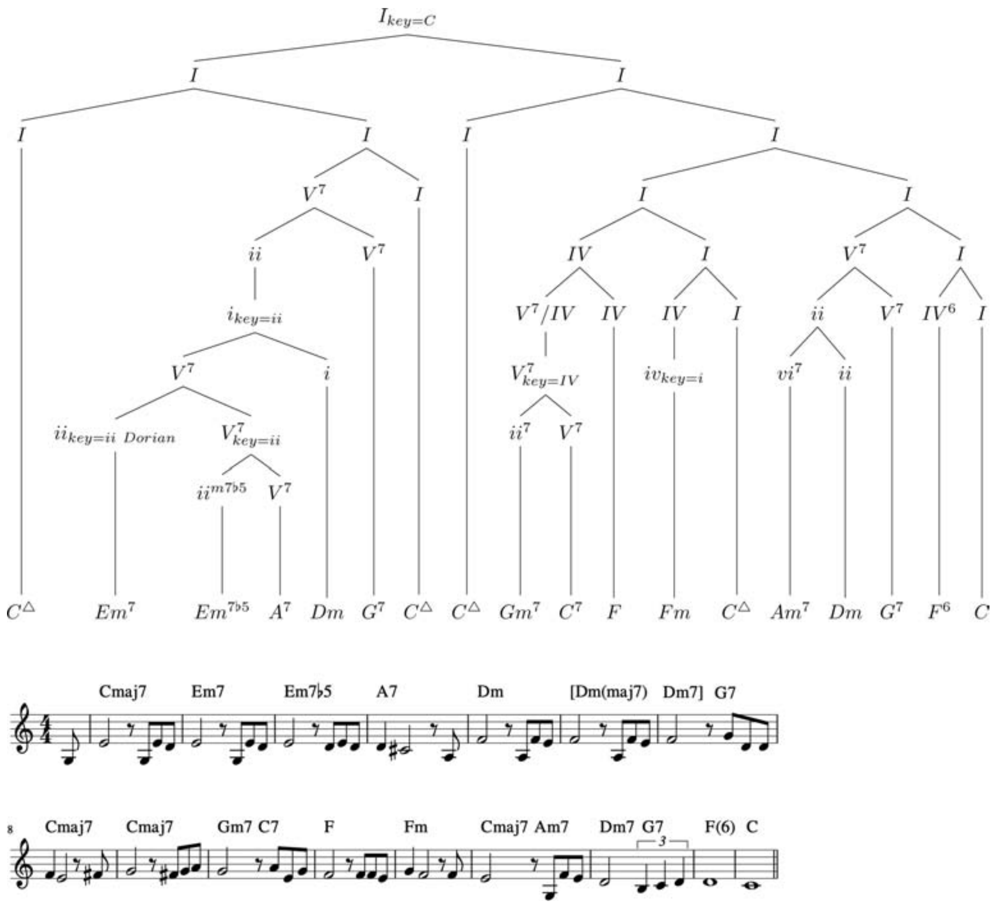
In the case of "My Way," the use of the subdominant throughout the entire phrase is so pronounced that relegating these events to the status of mere improper harmony appears implausible. The fact that the subdominant derivation from the tonic in (m. 10–11) can be so stable that it is productive in itself (and could also be in m. 15) provides strong evidence that the left-recursive subdominant derivation off the tonic features the same status as the dominant derivation. Therefore, this example may be considered a counterargument to Katz and Pesetsky's point for the need of internal merge based on progressions requiring a strict "lock."

This finding, however, does not necessarily affect the conclusion of the overall discussion concerning the formal complexity class of harmonic syntax in music, and

131 Katz and Pesetsky, "The Identity Thesis for Language and Music."

132 Nicole Biamonte, "Triadic Modal and Pentatonic Patterns in Rock Music," *Music Theory Spectrum*, no. 32 (2 2010): 95–110; Biamonte, "Triadic Modal and Pentatonic Patterns in Rock Music"; David Temperley, "The Cadential IV in Rock," *Music Theory Online* 17, no. 1 (2011).

Figure 22: An analysis of the first phrase of "My Way".



whether it lies in the context-free or mildly-context-sensitive class[133]—which is the main motivation behind Katz and Pesetsky's internal merge argument. Given the argument above, the dominant-tonic lock appears less plausible (and may also be modeled with an entirely context-free rule $I \rightarrow V' I'$, in which I' and V' are pre-surface symbols such that I' may not expand and V' may not expand to the right [in case, a theorist may assume the necessity of right-branching rules that are argued to be implausible in this article].)

However, one may still maintain the argument for the (mild) context-sensivity of harmony.[134] One kind of dependency relation which cannot be modeled using context-free

133 Weir, "Characterizing Mildly Context-Sensitive Grammar Formalisms"; Aravind K. Joshi, K. Vijay Shanker, and David Weir, *The Convergence of Mildly Context-Sensitive Grammar Formalisms*, technical report MS-CIS-90-01 (University of Pennsylvania, Department of Computer and Information Science, 1990), https://repository.upenn.edu/cis_reports/539.

134 Note that if the whole of music is concerned, repetition structure as in the context of musical form immediately

rules or internal merge only, and might thus provide evidence for mild context-sensitivity, is the half cadence that establishes dependencies that are, to some extent, similar to right-node raising in language.[135] Another case that may require internal merge is the modeling of double functions of chords (such as the same surface chord playing a role as an overarching tonic and a *V/IV* in two different subtrees[136], or the double function of the *ii–V–I* progression at the beginning of "Tune Up" discussed above). In this case, the lower function would internally merge with the adjacent derivation of the same surface chord from a higher branch in the tree with the result that the second instance is not "pronounced."[137] Finally, Naphtali Wagner provides an illuminating assessment of crossing branches in Schenkerian analysis in contrast with the GTTM that rules out such cases.[138] The incorporation of such cases into a theory at the level of polyphonic voice-leading would also require a mechanism beyond context-free complexity. All of these points, in addition to the trivial context-sensitivity at the level of form, speak strongly in favor of musical structure exhibiting complexity beyond context-freeness.

7. CONCLUSION

This article has proposed a formalization of Jazz harmony in terms of a formal, generative syntax based on a small set of core chord dependency relations that formulate a correspondence between harmonic structure and its interpretation based on the specific functional role of each chord in a sequence. The syntactic theory assumes that harmonic dependencies at local as well as large-scale levels may be derived from the two main principles of prolongation and preparation, as well as the orthogonal rule of substitution. It further proposed a formalization of the derivation of out-of-key chords and an understanding of modulation as the casting of a pivot chord and its recursively embedded elaboration under its own key scope. Because of its recursive nature, the proposed formalism models harmonic relations at different timescales, from local progressions to global harmonic

pushes the required complexity beyond context-freeness since structural repetition cannot be modeled with a context-free formalism.

135 Rohrmeier and Neuwirth, "Towards a Syntax of the Classical Cadence"; see also Granroth-Wilding and Steedman, "A Robust Parser-Interpreter for Jazz Chord Sequences."

136 For an example see, e.g., Rohrmeier, "Towards a Generative Syntax of Tonal Harmony."

137 I would like to thank Oriol Quintana Sanfeliu for pointing me to such an understanding of the double function derivation.

138 Naphtali Wagner, "No Crossing Branches? The Overlapping Technique in Schenkerian Analysis," *Theory and Practice* 20 (1995): 149–175.

relations. This also makes it possible to account for the harmonic and modulatory spine of common (Jazz) forms. The formalism captures tail-recursive dependencies as well as center-embedded harmonic structures, each of which frequently occurs in Jazz. As a theoretical framework the proposed syntax theory does not stand in opposition to theories of musical schemata or voice-leading; it rather offers a particular perspective to account for structural relations at the level of chord sequences.

The ideas put forward in this article constitute a step towards a new form of music theory that embraces impulses, results, and standards from closely related disciplines such as mathematics, computational modeling and music psychology.[139] As an advancement in this vein, the syntactic theory proposed in this article is sufficiently precisely formulated that it lends itself to computational implementation, evaluation, and further development.[140] After all, theories of musical structure are what Geraint Wiggins calls "pen and paper" computational models of music, and it is the natural next step to sharpen them and to make them computationally accessible. Reliable fine-grained development of subtle aspects of theories inevitably requires a computational toolkit. In this spirit this article aims to provide an example of musical theory building in the age of the digital revolution.

ACKNOWLEDGEMENTS

I am particularly grateful to Aki Takase for ever illuminating, creative and insightful conversations on the spirit of Jazz harmony. I owe many thanks to Daniel Harasim, Markus Neuwirth, Christoph Finkensiep, Jonah Katz, Andrew McLeod, Michael Polth, Fabian Moss, Thomas Noll, Philippe Schlenker, Mark Steedman, Robert Lieck, and Boyan Beronov for their comments and their ever ongoing, stimulating, and challenging discussions.

---

139 Anja Volk and Aline Honingh, "Mathematical and Computational Approaches to Music: Challenges in an Interdisciplinary Enterprise," *Journal of Mathematics and Music* 6, no. 2 (2012): 73–81; Anja Volk, Frans Wiering, and Peter Van Kranenburg, "Unfolding the Potential of Computational Musicology," in *Proceedings of the 13th International Conference on Informatics and Semiotics in Organisations* (2011); Guerino Mazzola, *Geometrie der Töne: Elemente der Mathematischen Musiktheorie* (Heidelberg: Springer-Verlag, 2013); Pearce and Rohrmeier, "Music Cognition and the Cognitive Sciences"; Wiggins, "Computational Models of Music"; Neuwirth and Rohrmeier, "Wie wissenschaftlich muss Musiktheorie sein? Chancen und Herausforderungen musikalischer Korpusforschung."

140 De Haas, "Music Information Retrieval Based on Tonal Harmony"; Harasim, Rohrmeier, and O'Donnell, "A Generalized Parsing Framework for Generative Models of Harmonic Syntax"; Harasim, O'Donnell, and Rohrmeier, "Harmonic Syntax in Time: Rhythm Improves Grammatical Models of Harmony."

FUNDING STATEMENT

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (GA No 760081–PMSB). I thank Claude Latour for supporting this research through the Latour Chair in Digital Musicology.

APPENDIX

The following paragraphs provide a brief introduction of the main ideas behind dependency graphs, trees, constituents and rewrite grammars for music theorists who may not be familiar with them. For precise and detailed formalizations see, for instance, the textbook by Michael Sipser.[141]

Dependency Structures

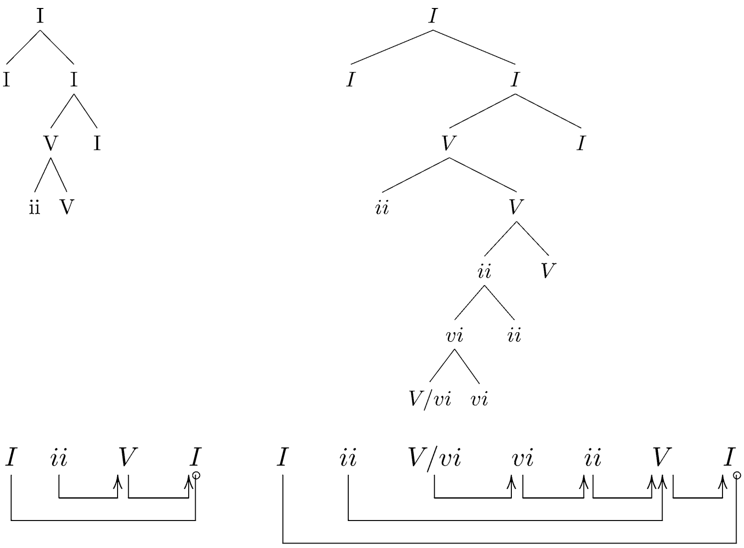
In functional harmonic sequences, chords express relations to other chords. For instance, a *I–ii–V–I* sequence may be analyzed such that *ii* prepares *V*, *V* prepares the final *I* and that the initial *I* and the final *I* form an overarching prolongation; crucially, the initial *I* does not imply or prepare *ii* in this sequence, or is connected to it in another way. A more complex example would be *I–ii–V/vi–vi–ii–V–I*. Here, *V/vi* prepares *vi*, *vi* prepares *ii*, both instances of *ii* prepare *V*, *V* prepares *I*, and the final *I* prolongs the initial *I*. This example exhibits two dependencies that do not refer to the immediate next chord (nonlocal dependencies), *I–I* and *ii–V*. Several theories of (Jazz) harmony denote dependency relations like in these examples with arrows between the chord symbols.[142] This kind of notation is formalized in a dependency graph (see the two bottom graphs in Figure 23). The two kinds of dependencies used here (preparation and prolongation) are drawn with different connectors. The same kinds of dependency relations are also expressed in the trees above the two dependency graphs. Each branch in the tree expresses one dependency relation. Crucially, every chord (except for the single root) establishes a dependency with another chord, and neither the dependency relations nor the branches in the trees cross each other. The types of dependency relations and grammars correspond to dependency grammars[143];

141 Sipser, *Introduction to the Theory of Computation*.

142 E.g., Mulholland and Hojnacki, *The Berklee Book of Jazz Harmony*; Chris Stover, "Jazz Harmony: A Progress Report," *Journal of Jazz Studies* 10, no. 2 (2014): 157–197. William Caplin notates applied dominants this way, Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*.

143 David G. Hays, "Dependency Theory: A Formalism and Some Observations," *Language* 40, no. 4 (1964): 511–525; Marco Kuhlmann, *Dependency Structures and Lexicalized Grammars: An Algebraic Approach* (Berlin-Heidelberg: Springer, 2010).

Figure 23: Expressing syntactic relations by a dependency graph and a syntax tree.



this also holds for GTTM's trees, a point that was raised by Christopher Longuet-Higgins in his insightful review of the GTTM[144]). However, because dependency graphs merely express relations between surface chords, relations such as substitution, complex functions (such as *V/V/V*) or modulation as formulated in this article cannot be expressed in such a simple way and require additional formal tools.

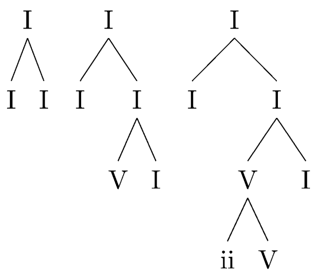
Constituents

Harmonic sequences may be the elaboration of a single overarching harmony (the *head*). For instance, when a tonic chord is to be extended/prolonged for two bars, one very common strategy is to achieve this by inserting dominants or a short stabilizing *ii–V–I* prolongation (such that the tonic prolongation becomes *I–V–I* or *I–ii–V–I*), or by using a plagal prolongation such as *I–IV–I*.[145] These relations can be expressed for this example using trees as shown in Figure 24. For each tree and subtree, the root indicates the *head* that is elaborated by its children (nodes or subtrees). A single chord or the group of chords that elaborates a single chord as the head forms a *constituent*.

Note that all the trees in the figure share the common root node *I*, while their children and leaves generate different extensions of *I* by combining prolongation and preparation

144 Christopher Longuet-Higgins, "All in Theory—The Analysis of Music," *Nature* 304, no. 5921 (1983): 93.
145 For a list of options of harmonic prolongation in the Classical style, see, e.g., Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*.

Figure 24: The figure shows the uses of syntactic structure to express overarching tonic prolongation using recursively embedded preparation. The trees show the similarity of the deep structure between the three sequences *I I*, *I V I* and *I ii V I*



in terms of binary branching. All trees here have in common that the final surface sequence produces a prolongation of the top node of the tree (*I*). Such prolongational insertions may result in non-local relations; the original *I–I* prolongation of the first tree is only implicit in the surface sequence of the final tree (while the tree analysis renders the *I–I* prolongation as the deepest relation that is first derived).

Grammar and Rules

The relations between chords that are expressed in the analyses above correspond with formal grammars and strings generated from the application of rules.

A grammar consists of surface (or *terminal*) symbols, *nonterminal* symbols (or *categories*), rewrite rules, and a start symbol. The example sequences in Figure 24 could be derived from the following set of rules: $I \rightarrow I\ I$, $I \rightarrow V\ I$, $V \rightarrow ii\ V$, and the start symbol *I*.

The rules of the grammar are applied recursively in order to generate complex sequences from simple ones via rewrite operations, beginning from the start symbol and successively applying the rules in the grammar one by one. In each rewrite operation, one rule of the set of available rules is applied to one of the symbols in the sequence and replacing it by the right-hand side of the arrow (→) in order to derive a more elaborate sequence. Because only one symbol is replaced at a time and irrespective of its surrounding context in the case of this type of (context-free) grammar, the order of rule applications does not matter. For instance, to derive the example sequence, starting with a single »*I*«, *I* is rewritten as »*I I*« (applying $I \rightarrow I\ I$), then as »*I V I*« (applying $I \rightarrow V\ I$), then as »*I ii V I*« (applying $V \rightarrow ii\ V$). Read from top to the bottom, the tree structure reflects the applications of rewrite rules that generate the analyzed sequence (see Figure 23).

The trees as used here have the property that the chord label of the parent node is identical to one of the children (as, e.g., in Figure 23). This comes from the music-specific property that the trees model the elaboration of a nonterminal harmonic event and indicates that, for cases in which the two children are different, one of the children is hierarchically more *stable* than the other(s) and can act as the *head* of the entire generation below.

Abstract

The regularities underlying the structure building of chord sequences, harmonic phrases, and combinations of phrases constitute a central research problem in music theory. This article proposes a formalization of Jazz harmony with a generative framework based on formal grammars, in which syntactic structure tightly corresponds with the functional interpretation of the sequence. It assumes that chords establish nested hierarchical dependencies that are characterized by two core types: preparation and prolongation. The approach expresses diatonic harmony, embedded modulation, borrowing, and substitution within a single grammatical framework. It is argued in the second part that the proposed framework models not only core phrase structure, but also relations between phrases and the syntactic structures underlying the main forms of Jazz standards. As a special case, the Blues form relies heavily on the plagal derivation from the tonic and is analyzed in comparison with other analytical approaches to the Blues. The proposed theory is specified to a sufficient level of detail that it lends itself to computational implementation and empirical exploration, and this way it makes a step towards music theory building that embraces the close links between formal, mathematical, and computational methods.

About the Author

Martin Rohrmeier is the director of the Digital and Cognitive Musicology Lab (DCML) at the École Polytechnique Fédéral de Lausanne (EPFL). He studied philosophy, mathematics, and musicology in Bonn, and received his master and doctoral degrees from Cambridge University at the Centre for Music & Science under the supervision of Ian Cross. He subsequently worked at Microsoft Research, FU Berlin, and MIT linguistics with a fellowship from the MIT Intelligence Initiative. In 2014 he was appointed one of the Open Topic Professorships in Systematic Musicology and Music Cognition at TU Dresden. In 2017 he joined the EPFL as associate professor for Digital Musicology.