REHARMONISATION GRAMMAR RULES IN JAZZ IMPROVISATION

George KONSTANTELAKIS (Γιώργος Κωνσταντελάκης) (gkwnstantelakis@gmail.com) [1] and
Christina ANAGNOSTOPOULOU (Χριστίνα Αναγνωστοπούλου) (chrisa@music.uoa.gr) [1]

[1] *Department of Music Studies*, **National and Kapodistrian University of Athens**, Athens, Greece

ABSTRACT

Reharmonisation is a core mechanism through which jazz improvisers shape cadential drive, local tension, and harmonic direction. This paper proposes an interpretable grammar of jazz reharmonisation consisting of 10 generative rules (19 including subrules), derived from a corpus of 30 piano improvisations using left-hand voicings as evidence of harmonic intent and informed by established accounts of harmonic function and jazz reharmonisation. The grammar formalizes recurrent mechanisms including quality reinterpretation, tritone substitution, cadential insertion and omission, diminished interpolation, and chromatic approach. Evaluation suggests that reharmonisation is widespread and frequently compositional. In the 30-improvisation source corpus, 50.5% of 192 annotated occurrences require multi-rule primary explanations, and 25.5% admit a second annotated explanation. In an external validation on 30 additional improvisations, the fixed rule inventory achieves full coverage of 350/350 under the present annotation protocol, with 60.6% of occurrences requiring multi-rule explanations and 26.0% admitting a second annotated explanation. These results suggest that the rule inventory transfers beyond the source corpus. Implemented in an explainable prototype ("Reharmonizer"), the transformation grammar generates alternative reharmonisations with explicit derivation traces. The study thus extends grammar-based harmonic modeling to improvisational reharmonisation and offers a compact rule inventory for analysis, computation, and pedagogy.

1. INTRODUCTION

Reharmonisation is a fundamental aspect of jazz improvisation [1], [2]. Beyond melodic variation, improvisers often reshape an underlying progression through dominant preparation, substitution, and cadential compression or expansion. Although these practices are commonly taught as techniques, they also recur in consistent local patterns that can be described explicitly.

This paper proposes a rule-based formalization of functional reharmonisation in improvisation: an inventory of 10 generative rules (19, with subrules). The rule set serves three connected purposes: as an analytical vocabulary for annotating and comparing reharmonisation events, as executable operators for generating and testing candidates, and as a pedagogical representation that links analytical labels to concrete chord-sequence rewrites through derivation traces. The rules are derived from corpus observation, jazz theory literature, and prior grammar-based work. Their operational value is examined through a computational prototype, developed as part of ongoing PhD work, that implements them as executable transformations. The main contribution is therefore a compact, corpus-grounded rule inventory that supports explanation, generation, and pedagogy through explainable harmonic rewrites.

Copyright: © 2026. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1.1 Related Work and Positioning

Grammar-based and hierarchical approaches, originating in formal linguistics [3], have long informed computational accounts of harmony. Steedman's generative description of jazz chord sequences [4] provides an early formal model, while Rohrmeier's work on tonal and jazz harmony [5], [6] develops grammar-based accounts of harmonic syntax and form. Related work has extended these ideas through parsing, probabilistic modeling, rhythm-aware representation, and corpus-oriented evaluation [7–9].

Early symbolic AI research also investigated explicit representations of jazz harmonic knowledge. Hirata proposed a deductive object-oriented framework for symbolic reasoning about jazz harmony [10], while subsequent work extended this approach to rule-based jazz chord reharmonisation [11].

Computational creativity research has also addressed jazz harmony through grammar-based generation, functional-harmonic modeling, and constraint-based systems [12–14]. Other systems treat reharmonisation-adjacent transformations without adopting an explicit grammar formalism, including learning-based jazzification, statistically constrained reharmonisation, and graph/path-based approaches to harmonic redirection [15–17].

Recent work has explored complementary approaches to reharmonisation. Moreira examines non-functional chord substitution that extend beyond traditional functional harmony [18, 19]. Christou and Dannenberg formulate reharmonisation as a sequence-to-sequence deep learning problem, moving from a diatonic vector space to a chromatic vector space [20]. Both these approaches capture transformations that bypass functional rules. The present work instead focuses on function-based reharmonisation, with ex-

plicit derivation paths for both analysis and computational generation.

Building on earlier grammar-based and symbolic AI research, the present work models performed reharmonisation as local transformation operations, providing an explicit, interpretable rule inventory grounded in improvisational practice. The main contribution is an interpretable, corpus-grounded transformation grammar for reharmonisation, together with evidence of its analytical adequacy and computational realizability.

2. METHODOLOGY

2.1 Corpus Selection and Analysis

The rule inventory was developed through a hybrid inductive–deductive process: core mechanisms were abstracted from patterns in a 30-item corpus and completed using established principles of functional jazz harmony [21–29]. The corpus focuses mainly on jazz pianists from the 1950s to the present and spans stylistic traditions within functional harmony; the aim is musical diversity rather than statistical representativeness. Piano improvisations were selected in order to use left-hand voicings as evidence of harmonic intent, reducing the ambiguity of monophonic lead-line interpretation. The 30 improvisations and analyzed excerpt entry points are listed in Supplementary Table S1. The improvisations vary in duration, but each contains at least one full chorus over the theme.

Each excerpt was annotated measure by measure, or event by event where necessary, using bar/beat/sub-beat indices in a structured sheet (Supplement S2). A reference progression was defined from a lead-sheet or standard-changes version, minimally adjusted only when verified audio implied a different harmonic rhythm. A reharmonisation occurrence is defined as a continuous span in which the performed harmony diverged from this reference until the next alignment point. Published or online transcriptions were used where available and manually verified against the recordings. Candidate cases were retained only when the divergence was better explained as harmonic reinterpretation than as melodic chromatic embellishment, based on left-hand support and guide-tone stability.

The annotation process consists of a sequence of analytical stages: (1) referencing lead sheet and improvised transcription, (2) grouping performed notes per chord/measure, (3) identifying and removing non-harmonic tones, (4) extracting harmonic pitch content, (5) assigning the performed chord label, (6) comparing with the reference harmony to detect reharmonisation, and (7) deriving the observed transformation using the proposed grammar. A complete worked example illustrating each stage is provided in the supplementary material.

2.2 Rule Derivation

From the collected observations, recurrent harmonic transformations were abstracted into generalized rewrite rules of the form X→Y. The inventory aims to capture both descriptive recurrence and generative potential: some rules reflect frequent corpus patterns directly, while others generalize from observed tendencies to cover musically plausible but less common reharmonisations. Because functional jazz improvisation admits open-ended local transformations, some rules are intentionally broad, with contextual preference handled by the annotation protocol and, computationally, by ranked derivation choice. The rules target harmonic function rather than strict durational equivalence. A rewrite may expand, compress, sustain, or redistribute a local span, consistent with rhythm-aware approaches to harmonic syntax [9]. The inventory overlaps only partially with earlier grammar-based models [4–6], because it is designed for performed reharmonisation rather than a standard's underlying harmonic syntax. The next section presents ten rules and nine subrules.

3. GRAMMAR RULES OF REHARMONISATION

3.1 Rule Presentation

The grammar rules are presented below in transformation notation, each accompanied by a brief explanation. They are grouped conceptually into quality change (sub), substitution (sub), prolongation/preparation (prol), omission/compression (del/comp), added-chord expansion (ins), and chromatic approach (ins). For presentation purposes, the rules are also grouped by relation to prior grammar-based work: closest to Steedman, closest to Rohrmeier, proposed here as extensions or adaptations, and proposed here without a close direct precedent.

Progressions are modeled as ordered sequences of chord symbols. In the schemata, X denotes a pitch-class root, "q" represents a chord quality, and root arithmetic is expressed in semitone units mod 12; thus X+k indicates transposition by k semitones, and tritone-related roots are written as Xsub = X+6. Juxtaposition denotes chord concatenation, parentheses indicate optional sustain, and "/" separates alternative rewrites. Roman numerals are interpreted relative to the local tonic center of each occurrence.

3.1.1 Steedman-derived rules

R3. Prolongation (PROL)
Schema: X → X X
Description: A chord or harmonic unit is prolonged through repetition.
Constraint: this rule is usually the first step, after which a second rule is applied to X1 or X2 of schema.

R7a. Preparation of an Existing ii–V by the Dominant of its Dominant (INS)
Schema: ii V → V/V ii V
Description: Preparation of an existing ii–V by inserting the dominant of its dominant (adds one preparatory step before the unit).

R9. Diminished-Chord Interpolation and Symmetric Expansion (INS (+ SUB equivalence))
Schema: X → X°7 X (or Y°7 Xq)
Description: Diminished chords may be interpolated before a target harmony and can also generate families of related dominant or diminished structures through symmetric transposition.

3.1.2 Rohrmeier-derived rules

R2. Tritone Substitution (SUB)
Schema: X7 → Xsub7, where Xsub = X + 6 (mod 12)
Description: A dominant chord may be replaced by the dominant rooted a tritone away.
R7. Preparation with Cadence (INS)
Schema: X → (ii–V of X) X / (V of X) X
Description: A target chord is prepared by insertion of its dominant or a full ii–V cadence.
R8. Addition of ii before a Dominant (INS)
Schema: X7 → ii X7
Description: A dominant chord can be expanded backward into a ii–V by inserting its corresponding supertonic.

3.1.3 Rules proposed here as extensions

R2a. Tritone Substitution of a ii–V Cadence (SUB)
Schema: ii V → ii_sub V_sub
Description: Tritone substitution applied to a complete ii–V cadence.
R8a. Minor Chord Reinterpreted as ii of a New Local Cadence (SUB+INS)
Schema: Xm → ii V
Description: A minor sonority may be reinterpreted as the ii of a new local cadence, yielding a generated ii–V trajectory.
R9a. Diminished Minor-third Cycle (equivalence class) (SUB)
Schema: X°7 → (X+3)°7 / (X+6)°7 / (X+9)°7
Description: Minor-third cycle of equivalent diminished structures (same pitch-class collection under transposition).
R9b. Diminished Major-third Expansion (rarer) (SUB)
Schema: X°7 → (X+4)°7 / (X+8)°7
Description: Major-third symmetric expansion used in more distant color shifts.
R10. Chordal Chromatic Approach (INS)
Schema: X → A X, where root(A) ∈ {root(X) - 1, root(X) - 2} (mod 12)
Description: A chord may be approached by a chromatic harmony from one or two semitones below. Many realizations can be interpreted via combinations of secondary dominants, tritone substitution/backdoor logic, and temporary functional reinterpretation; the rule groups them as a compact approach family.
Constraint: A is an approach chord whose quality is context-dependent.

3.1.4 Rules proposed here without a close direct precedent

R1. Quality Change to Dominant (SUB)
Schema: Xq → X7
Description: Any chord quality may be reinterpreted as a same-root dominant to increase directional motion toward the following harmony. It is especially common in cadential contexts, but may also occur in looser local rewritings. Its plausibility is therefore context-dependent.
R1a. Dominant → Minor or Half-diminished (SUB)
Schema: X7 → Xm7 or X∅7
Description: A dominant chord may be reharmonised as a minor or a half-diminished sonority on the same root, more usually in cases where it remains embedded in a dominant-to-dominant chain.
R1b. Minor/Major Quality Alternation (SUB)
Schema: Xm7 ↔ Xmaj7
Description: Rare quality alternation between minor and major seventh sonorities on the same root, typically as local color (modal mixture / tonic-color shift).
R2b. Tritone Shift with Preserved Quality (SUB)
Schema: Xq → Xsubq
Description: Generalized tritone substitution with preserved chord quality outside dominant contexts (rare).
R4. Chord Omission / Temporal Absorption (DEL/COMP)
Schema: Xq Yq → Xq Xq / Yq Yq
Description: Because of rhythmic displacement and temporal compression, one harmony in a local pair may be omitted in performance, so that either or occupies the span. The rule is intentionally general and captures temporal absorption as a recurrent improvisational mechanism whose plausibility depends on context.
R5. Omission of ii in a ii–V (DEL/COMP)
Schema: ii V → V (V)
Description: In fast tempos or dense harmonic contexts, improvisers often omit the ii chord and retain only the dominant across the cadential span.
R5a. Omission of V in a ii–V (DEL/COMP)
Schema: ii V → ii (ii)
Description: Symmetric case in which the dominant is omitted and the ii is sustained across the span.
R6. Full Cadence Omission (Tonic Prolongation) (DEL/COMP)
Schema: ii V I → I I I
Description: A complete cadence may be abstracted by sustaining the tonic area.

3.2 Rule Summary Table

Layer / ID	Rewrite type	Transformation schema (canonical)
Rules with closest precedent in Steedman		
R3	PROL	X → X X
R7a	INS	ii V → (V/V) ii V
R9	INS	X → X°7 X (or Y°7 Xq)
Rules with closest precedent in Rohrmeier		
R2	SUB	X7 → Xsub7 (Xsub = X+6 mod 12)
R7	INS	X → (V of X) X / (ii–V of X) X
R8	INS	X7 → ii X7
Rules proposed here as extensions/adaptations		
R2a	SUB	ii V → ii_sub V_sub
R8a	INS	Xm → ii V
R9a	SUB	X°7 → (X+3)°7 / (X+6)°7 / (X+9)°7
R9b	SUB	X°7 → (X+4)°7 / (X+8)°7
R10	INS	X → A X; root(A)∈{root(X)-1, root(X)-2}
Rules proposed here without a close direct precedent		
R1	SUB	Xq → X7
R1a	SUB	X7 → Xm7 or X∅7
R1b	SUB	Xm7 ↔ Xmaj7
R2b	SUB	Xq → Xsubq
R4	DEL/COMP	X Y → X X / Y Y
R5	DEL/COMP	ii V → V (V)
R5a	DEL/COMP	ii V → ii (ii)
R6	DEL/COMP	ii V I → I I I

Table 1. Rule summary

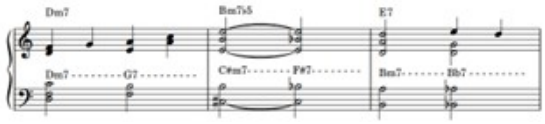


Figure 1. The Days of Wine and Roses, McCoy Tyner. Rule-path: R8a → R2 → R7 (cadential expansion, tritone substitution, preparation).

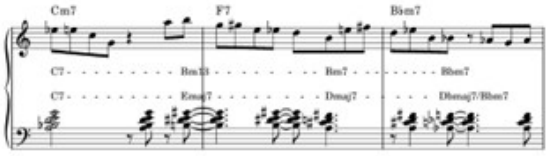


Figure 2. Not You Again, Brad Mehldau. Rule path i: R1 → R2 → R1a (quality change, tritone substitution). Rule path ii: R1 → R7 → R2 → R1a/R1b, R9 (quality change, preparation, tritone substitution, quality change, diminished).

3.3 Analytical Examples

Figure 1 shows a reharmonisation by McCoy Tyner in 'The Days of Wine and Roses'. The underlying segment Dm7 — Bm7b5 — E7 is reharmonised as Dm7 G7 — C#m7 F#7 — Bm7 Bb7 —. One plausible rule-based reading is: (R8a) expands Dm7 by adding G7; then (R2) applies tritone substitution to E7 (yielding Bb7); and (R7) prepares Bm7 with its ii-V (C#m7-F#7). This example illustrates how a short passage can combine functional preparation and substitution within a compact derivation.

Figure 2 shows a broader source of ambiguity in the annotation process: the performed material can support more than one musically acceptable chord-symbol labeling for the same reharmonised passage. The performed notes clearly diverge from the original harmony, establishing that a reharmonisation has occurred; the question is how to label it. For instance, the voicing A–C#–D–F# at measure 2 admits both a Bm9 (as a 7-9-3-5 minor voicing and a Dmaj7 reading, since the two chords share identical pitch-class content. A first path starts with Cm → C7 (R1), then treats F7 → B7 via tritone substitution (R2), and finally reinterprets B7 as Bm through reverse quality reinterpretation (R1a). An alternative path targets a Bbm/Dbmaj area: a local preparation (R7) introduces dominant material Ab7, tritone substitution (R2) redirects the cadence to D7, and quality reinterpretation (R1a)/(R1b) supports the shift in local function, Dmaj. The Emaj7 event may be interpreted through the diminished/symmetric family (R9)/(R9a), highlighting how the framework can document analytical ambiguity rather than forcing a single reading.

4. ANALYTICAL AND COMPUTATIONAL USE OF THE RULE SYSTEM

4.1 The Rules as an Analytical Framework

Even without computational implementation, the rules provide a practical framework for describing reharmonisation in improvisation. Instead of treating departures from a lead-sheet progression as isolated chord-symbol differences, the analysis represents them as specific transformation types such as reinterpretation, substitution, cadential preparation, omission, and chromatic approach. This makes comparison easier across solos and performers, including cases that differ on the surface but fulfill similar local functions. The framework also captures the fact that many reharmonisations are multi-step. A single local event may require a sequence of rule applications rather than a single substitution label. When more than one derivation is musically plausible, the present annotation records a preferred reading under the current protocol while allowing alternatives to remain visible.

4.2 Computational Prototype: Rule-Based Reharmonisation System

We implemented a symbolic reharmonisation prototype ("Reharmonizer") that maps an input chord progression to a page of alternative progressions under user controls, including target output length, optional locks/constraints (e.g., tonic center or fixed chords), and a continuous complexity control (0–100). Generation proceeds by bounded derivation search: candidates are expanded up to a maximum derivation depth and a maximum output length, and a rule is applied only when its rule-context predicate holds, as defined for each operator. Default search bounds and scoring settings used in the present experiments are listed in Supplementary appendix S4.

To manage combinatorics, mechanisms are grouped into coarse method families (F0–F6) used for scheduling, gating, and diversity control. The rules discussed here are implemented as executable local-edit operators within these families (primarily F0–F4), so that each generated output remains explainable as a sequence of substitutions, insertions, prolongations, and absorptions with an explicit derivation trace (rule IDs plus intermediate progressions). Candidates are filtered by hard constraints and ranked by a weighted objective favoring compactness, variety, and harmonic plausibility. After deduplication and novelty gating, the system returns the top-ranked candidates together with their derivation traces.

In the present paper, the prototype serves as evidence of computational implementability rather than as a fully evaluated standalone system. Its role is to show that the proposed rule-based system is explicit enough to support both analytical description and generative exploration, and to function as a consistency check on the formalism by requiring explicit rewrite operations, context predicates, and derivation-selection procedures.

The example below shows that given Fm — Db7 — G7 — Cmaj, the system may return the traced derivation R4 → R7, yielding Fm — Ebm7 — Ab7 — Db7 — Cmaj — Cmaj and thus representing reharmonisation as a sequence of interpretable operations rather than as a single substitution. Musically, the result both sustains the resolution harmony and strengthens the approach to Db7 with an ii–V preparation.

Reharmonizer Output:

```
Step 1:  R4 Dom→ResMaj
```

```
family:  F1 --- Functional diatonic
Edits, change:  replace G7 → Cmaj

before:  Fm | Db7 | G7 | Cmaj

after:  Fm | Db7 | Cmaj | Cmaj

Step 2:  R7 AddII-V@1

family:  F1 --- Functional diatonic,
approach:  ii--V insertion, change:
insert Ebm7, Ab7

before:  Fm | Db7 | Cmaj | Cmaj

after:  Fm | Ebm7 | Ab7 | Db7 | Cmaj |
Cmaj
```

5. EVALUATION

5.1 In-corpus validation

Both the in-corpus and external evaluation reflect author-annotated coverage under the present protocol. We first assessed the in-corpus adequacy of the reharmonisation grammar on 30 jazz improvisations, comprising 192 annotated reharmonisation occurrences. All 192 occurrences admit at least one explanation in the rule inventory (coverage = 192/192). Since this corpus also served in the construction of the rule inventory, the in-corpus evaluation should be understood as a test of internal adequacy rather than out-of-sample generalization. Accordingly, its main significance lies not only in full coverage under the present annotation protocol, but also in the extent to which many occurrences require compositional, multi-rule explanation. Table 2 shows that reharmonisation is widespread but uneven across performances (25/30 improvisations contain at least one occurrence; range 1–30 occurrences per active improvisation). A key result is the prevalence of composite derivations: 50.5% of primary explanations require two or more rule IDs, with a mean of 1.71 distinct rule IDs per occurrence. In addition, 25.5% of occurrences admit a second annotated explanation ("V2"), yielding 241 total explanation paths, of which 60.2% are multi-rule. Across primary and secondary explanations, all 19 rule IDs are activated; in the primary layer alone, 18/19 appear (unused: R9b). These results support a grammar-based account of reharmonisation over a single-label taxonomy. Annotation details and corpus metadata are provided in Supplementary Tables and Appendix S1–S2; Section 5.2 complements this internal view with external validation on a separate set of unseen improvisations.

Metric	Result
Annotated reharmonisation occurrences (N)	192
Improvisations in corpus	30
Improvisations with $\geq$1 occurrence ("active")	25 / 30 (83.3%)
Occurrences per active improvisation	mean 7.7, median 5, range 1–30
Multi-rule primary explanations	97 / 192 (50.5%)
Occurrences with a second explanation (V2)	49 / 192 (25.5%)
Total explanation paths	241
Rule inventory actually used	19 / 19 rules across primary+V2
Top 3 rules (share of primary occurrences)	R1 36.5%, R2 26.6%, R7 20.8%
Most frequent rule-pair (primary co-occurrence)	R1+R2 = 29 / 192 (15.1%)

Table 2. Corpus-level summary (N = 192 annotated occurrences; 30 improvisations): reharmonisation is widespread but uneven across performances, 50.5% of primary explanations are multi-rule, and 25.5% admit a second explanation.

Rule ID	Mechanism	Count	% of occurrences
R1	Quality/function reinterpretation	70	36.5%
R2	Tritone substitution	51	26.6%
R7	Added dominant / cadence preparation	40	20.8%
R4	Omission / absorption	33	17.2%
R5	ii–V omission / compression	23	12.0%
R5a	ii–V omission (variant)	20	10.4%
R8	Added ii before dominant	18	9.4%
R1a	Reverse quality reinterpretation	12	6.3%
R3	Prolongation	12	6.3%
R8a	Cadence from minor chord (Xm→ii–V)	12	6.3%
R10	Chromatic approach	11	5.7%
R1b	Quality alternation	6	3.1%
R2a	Tritone substitution (ii–V unit)	5	2.6%
R6	Cadential omission / tonic prolongation	4	2.1%
R9	Diminished interpolation	4	2.1%
R2b	Tritone shift with preserved quality	3	1.6%
R9a	Diminished symmetry (minor-third cycle)	3	1.6%
R7a	Preparation of an existing cadence	1	0.5%

Table 3. Rule usage distribution across the annotated corpus (non-exclusive counts under primary explanations)

5.2 External validation

To test out-of-corpus transferability, the reharmonisation transformation grammar was applied to a separate external set of 30 improvisations in which 350 reharmonisation occurrences were detected. Each occurrence was documented as an original-harmony → reharmonised-harmony pair together with the rule IDs required for the transformation. Under the present annotation protocol, all 350 occurrences are covered by the rule inventory (coverage = 100.0% under the present annotation protocol), and all 30 improvisations contain at least one reharmonisation event. The 30-item external validation set is listed in Supplementary Table S3. Importantly, the rule inventory was fixed before annotation of the external set, and no new rule was introduced based on these unseen improvisations. The external result should therefore be read as full author-annotated coverage on unseen material under the present protocol.

Metric	Result
Annotated reharmonisation occurrences (N)	350
Improvisations in external corpus	30
Improvisations with $\geq$1 occurrence ("active")	30 / 30 (100.0%)
Occurrences per improvisation	mean 11.7, median 10.5, range 3–23
Coverage by rule inventory	350 / 350 (100.0%)
Single-rule primary explanations	138 / 350 (39.4%)
Multi-rule primary explanations	212 / 350 (60.6%)
Mean distinct rule IDs per primary explanation	2.04
Primary derivation depth	1 rule: 39.4% ; 2 rules: 33.7% ; 3+ rules: 26.9%
Occurrences with a second explanation (V2)	91 / 350 (26.0%)
Total explanation paths	441
Mean explanations per occurrence	1.26
Single-rule explanation paths	147 / 441 (33.3%)
Multi-rule explanation paths	294 / 441 (66.7%)
Mean distinct rule IDs per explanation path	2.10
Rule inventory actually used	19 / 19
Top 3 rules, all explanation paths	R7 42.0%, R2 35.6%, R1 29.5%
Most frequent pair, all explanation paths	R2+R7 = 77 / 441 (17.5%)

Table 4. External corpus-level summary (N = 350 occurrences; 30 improvisations)

Overall, the external results provide stronger evidence that the frozen rule inventory transfers beyond the 30-item source corpus. At the same time, the external profile is not identical to the in-corpus results: it places greater weight on cadential preparation and tritone logic, shows a higher proportion of multi-step derivations, and makes analytical ambiguity empirically visible through the V2 annotations. This supports the view that the grammar generalizes across performers and tunes while remaining sensitive to corpus-specific stylistic emphases.

Rule ID	Mechanism	Count	% of explanation paths
R7	Added dominant / cadence preparation	185	42.0%
R2	Tritone substitution	157	35.6%
R1	Quality/function reinterpretation	130	29.5%
R4	Omission / absorption	111	25.2%
R3	Prolongation	75	17.0%
R5	ii–V omission / compression	70	15.9%
R1a	Reverse quality reinterpretation	60	13.6%
R5a	ii–V omission (variant)	39	8.8%
R1b	Quality alternation	30	6.8%
R8	Added ii before dominant	23	5.2%
R9a	Diminished symmetry (minor-third cycle)	18	4.1%
R9	Diminished interpolation	16	3.6%
R9b	Diminished major-third expansion	11	2.5%
R10	Chromatic approach	9	2.0%
R8a	Cadence from minor chord (Xm$\rightarrow$ii–V)	8	1.8%
R6	Cadential omission / tonic prolongation	6	1.4%
Grouped $<$1.0%	R2b, R7a, R2a	11	2.5%

Table 5. Rule usage distribution in the external corpus (non-exclusive counts)

6. DISCUSSION

The evaluation indicates that reharmonisation is often compositional rather than single-step. In the source corpus, 50.5% of primary explanations are multi-rule, and 25.5% of occurrences admit a second annotated explanation; across all 241 in-corpus explanation paths, 60.2% are multi-rule. In the external set, 60.6% of occurrences require multi-rule primary explanations, and 26.0% of occurrences admit a second annotated explanation; across all 441 external explanation paths, 66.7% are multi-rule. Taken together, these results support modeling reharmonisation as a grammar of combinable transformations rather than as a flat substitution taxonomy. The two datasets are not identical in profile. The source corpus gives greater weight to quality/function reinterpretation, while the external set foregrounds cadential preparation and tritone logic more strongly. Even so, both corpora point to the same broader conclusion: functional jazz reharmonisation is structured, recurrent, and often explainable through combinations of local operations rather than single isolated moves.

From a computational musicology perspective, the value of the rule inventory lies in making analytical categories executable. This allows a proposed explanation to be tested as an actual transformation path and allows corpora to be compared not only by chord labels, but by the mechanisms used to move from one harmony to another. In this sense, the prototype serves both as a generative tool and as a consistency check on formalism. The approach aligns with broader goals in computational musicology by making analytical categories explicit, testable, and comparable across corpora.

Pedagogically, the rule system is useful because it represents reharmonisation as a sequence of named steps rather than only as a final chord string. Students can therefore compare alternative preparations, substitutions, and omissions, and see how different harmonic outcomes can arise from the same starting progression.

The current study has several limitations. Annotation is performed by the authors and does not yet include inter-annotator validation. The corpus is designed for diversity rather than statistical representativeness.

A related concern is whether full coverage reflects genuine explanatory adequacy or an unfalsifiable rule set. Because the grammar targets tonal, function-based reharmonisation, its high coverage reflects this scope rather than an ability to derive arbitrary progressions: it is not designed to explain polytonal/bitonal voicings without a single functional root, non-functional modal reharmonisation, twelve-tone procedures, or purely aesthetic (e.g., timbral) choices. Where such material occurs, the grammar may still return a formally valid but musically unintended derivation. Corpus selection accordingly emphasized earlier stylistic periods, before non-tonal and sonority-based improvisation became more prevalent; as an example, Kenny Kirkland's improvisation on Black Codes (1985, Wynton Marsalis "Black Codes (From the Underground)") was excluded as an improvisation driven by sonority and non-tonal-harmonic elements rather than functional harmony. This shows coverage is bounded by harmonic domain, not by the specific corpus selected — a boundary the independent external validation (Section 5.2) does not contradict.

7. CONCLUSION

We proposed ten corpus-grounded reharmonisation rules (19 with subrules) for jazz improvisation and implemented them as executable operators with derivation

traces. Across the source corpus and the external validation set, the results show that reharmonisation is widespread and often requires multiple rule applications, supporting a grammar-based account over a single-label taxonomy. The main empirical contribution is evidence for compositionality, as many cases are better explained by combinations of rules than by a single substitution label. Future work will expand the dataset, evaluate inter-annotator agreement, compare the framework with other analytical approaches, and develop style-specific variants of the rule set. A further methodological step is to generalize the current primary/secondary annotation scheme into set-valued annotation, so that multiple musically valid derivations can be represented more systematically and compared across annotators. Another priority is to integrate voice-leading and phrase-structure constraints so that the model more closely reflects reharmonisation in performance. The proposed grammar provides a compact, interpretable framework for analyzing and generating reharmonisation, with evidence for compositional structure and cross-corpus transferability.

Supplementary appendix and supporting materials are available at: https://doi.org/10.5281/zenodo.21373487

8. REFERENCES

[1] P. Berliner, *Thinking in Jazz: The Infinite Art of Improvisation*. University of Chicago Press, 1994.

[2] H. Crook, *How to Improvise: An Approach to Practicing Improvisation*. Advance Music, 1991.

[3] N. Chomsky, "Three models for the description of language," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 113–124, 1956.

[4] M. J. Steedman, "A generative grammar for jazz chord sequences," *Music Perception*, vol. 2, no. 1, pp. 52–77, 1984.

[5] M. Rohrmeier, "Towards a generative syntax of tonal harmony," *Journal of Mathematics and Music*, vol. 5, no. 1, pp. 35–53, 2011.

[6] ——, "The syntax of jazz harmony: Diatonic tonality, phrase structure, and form," *Music Theory and Analysis*, vol. 7, no. 1, pp. 1–63, 2020.

[7] D. Harasim, M. Rohrmeier, and T. J. O'Donnell, "A generalized parsing framework for generative models of harmonic syntax," in *Proceedings of the International Society for Music Information Retrieval Conference*, 2018.

[8] D. Harasim, C. Finkensiep, P. Ericson, T. J. O'Donnell, and M. Rohrmeier, "The jazz harmony treebank," in *Proceedings of the International Society for Music Information Retrieval Conference*, 2020.

[9] D. Harasim, T. J. O'Donnell, and M. Rohrmeier, "Harmonic syntax in time: Rhythm improves grammatical models of harmony," in *Proceedings of the International Society for Music Information Retrieval Conference*, 2019.

[10] K. Hirata, "Towards formalizing jazz piano knowledge with a deductive object-oriented approach," in *Proceedings of the IJCAI'95 Workshop on Music and AI*, Montreal, Canada, 1995, pp. 77–80.

[11] K. Hirata and M. Goto, "Herbie-kun: A jazz chord reharmonizer based on a deductive object-oriented framework," Technical Report of IPSJ SIG Music and Computer (SIGMUS), Tech. Rep. 86 (1996-MUS-017), 1996.

[12] J. Gillick, K. Tang, and R. M. Keller, "Learning jazz grammars," in *Proceedings of the Sound and Music Computing Conference*, 2009.

[13] A. Choi, "Jazz harmonic analysis as optimal tonality segmentation," *Computer Music Journal*, vol. 35, no. 2, pp. 49–66, 2011.

[14] F. Pachet and P. Roy, "Markov constraints: Steerable generation of markov sequences," *Constraints*, vol. 16, no. 2, pp. 148–172, 2011.

[15] T.-P. Chen, S. Fukayama, M. Goto, and L. Su, "Chord jazzification: Learning jazz interpretations of chord symbols," in *Proceedings of the International Society for Music Information Retrieval Conference*, 2020.

[16] M. Kaliakatsos-Papakostas, K. Velenis, L. Pasias, C. Alexandraki, and E. Cambouropoulos, "An HMM-based approach for cross-harmonization of jazz standards," *Applied Sciences*, vol. 13, no. 3, p. 1338, 2023.

[17] H. Yamamoto, Y. Uehara, and S. Tojo, "Jazz harmony analysis with ε-transition and cadential shortcut," in *Proceedings of the Sound and Music Computing Conference*, 2020, pp. 316–322.

[18] T. Moreira, *Harmonic Explorations: A Journey into Advanced Jazz Harmony*. Haaheim Media Publications, 2023.

[19] ——, "Melodic treatment: Non-functional chord substitution in commercial music," in *Proceedings of the 11th International Conference on New Music Concepts*, Treviso, Italy, 2024.

[20] A. Christou and R. Dannenberg, "Automatic creation of chordal and melodic chromaticism," in *Proceedings of the 11th International Conference on New Music Concepts*, Treviso, Italy, 2024.

[21] R. Rawlins and N. E. Bahha, *Jazzology: The Encyclopedia of Jazz Theory for All Musicians*. Hal Leonard, 2005.

[22] D. Berkman, *The Jazz Harmony Book*. Sher Music Co., 2013.

[23] R. Felts, *Reharmonization Techniques*. Berklee Press / Hal Leonard, 2002.

[24] B. Ligon, *Jazz Theory Resources, Vol. 2*. Hal Leonard, 2001.

[25] D. Liebman, *A Chromatic Approach to Jazz Harmony and Melody*. Advance Music, 2015.

[26] T. Pease and K. Pullig, *Modern Jazz Voicings*. Berklee Press / Hal Leonard, 2001.

[27] B. Boyd, *Jazz Chord Progressions*. Hal Leonard, 1997.

[28] M. Levine, *The Jazz Theory Book*. Sher Music, 1995.

[29] J. Mulholland and T. Hojnacki, *The Berklee Book of Jazz Harmony*. Berklee Press, 2013.