

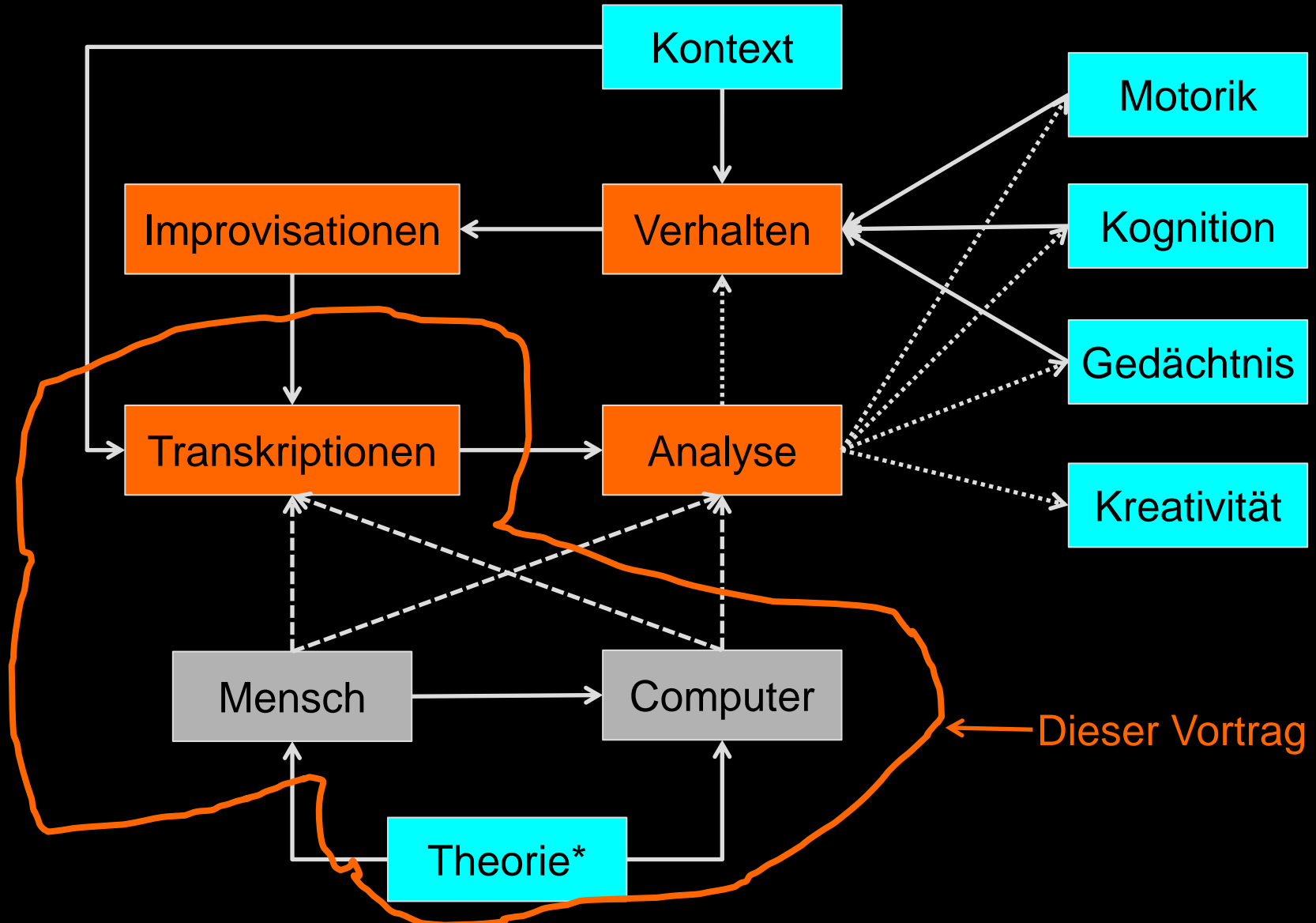
# LEBEN IN DER SCHNITTSTELLE

COMPUTERANWENDUNGEN UND MUSIKPSYCHOLOGIE  
IM KONTEXT DES JAZZOMAT RESEARCH PROJECT

Klaus Frieler, Martin Pfeiderer,  
Wolf-Georg Zaddach, Jakob Abeßer  
Institut für Musikwissenschaft, Weimar-Jena



# DAS JAZZOMAT-PROJEKT



# TRANSKRIPTIONEN: THEORIE

- Wie können valide Transkriptionen erstellt werden?
- **Ziel:** Performanz-orientierter statt notenbasierter Ansatz.
- **Lösung:** Sequenzen von Tonereignissen, die mit metrischen, harmonischen u.a. Kontextinformationen annotiert werden.
- **Annahme:** Tonereignisse mit definiertem Onset, definierter Tonhöhe, definierter Dauer:

$$e_n = (t_n, p_n, d_n)$$

# TRANSKRIPTION: THEORIE

- Als psychologische Größen unscharf und nur bedingt abbildbar auf physikalische Größen.
- Diese Beschreibung von Improvisationen scheint aber hinreichend für die meisten Jazzstile.
- Für andere eher insuffizient...



Chet Baker  
Long Ago and Far Away



Evan Parker  
Improvisation #1

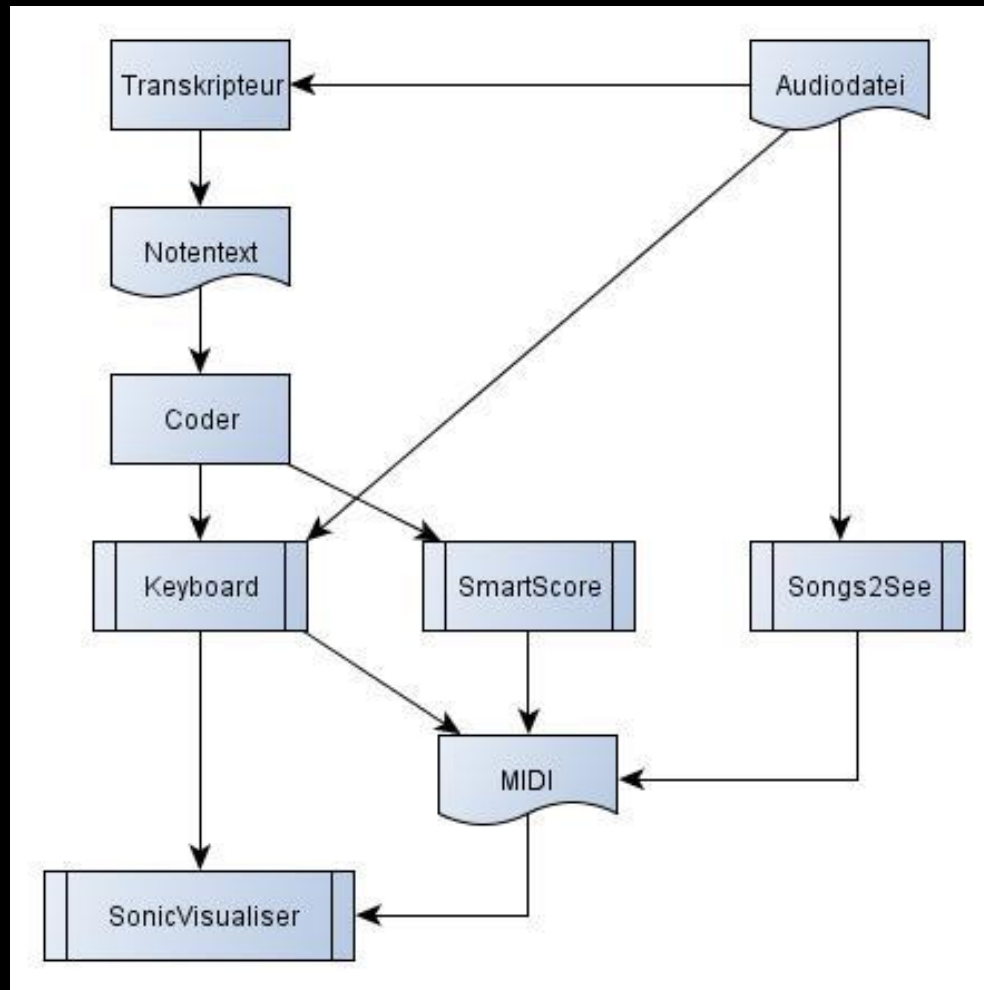
# TRANSKRIPTION: THEORIE

- Diese Beschreibung ist noch zu „dünn“ für viele interessante Analysen.
- Hinzukommen:
  - Metadaten,
  - harmonischer Kontext,
  - metrischer Kontext,
  - Phrasen,
  - Artikulationen,
  - Intensitäten.

# VON DER THEORIE ZUR PRAXIS...



# TRANSKRIPTION: PRAXIS



# TRANSKRIPTION: OMR

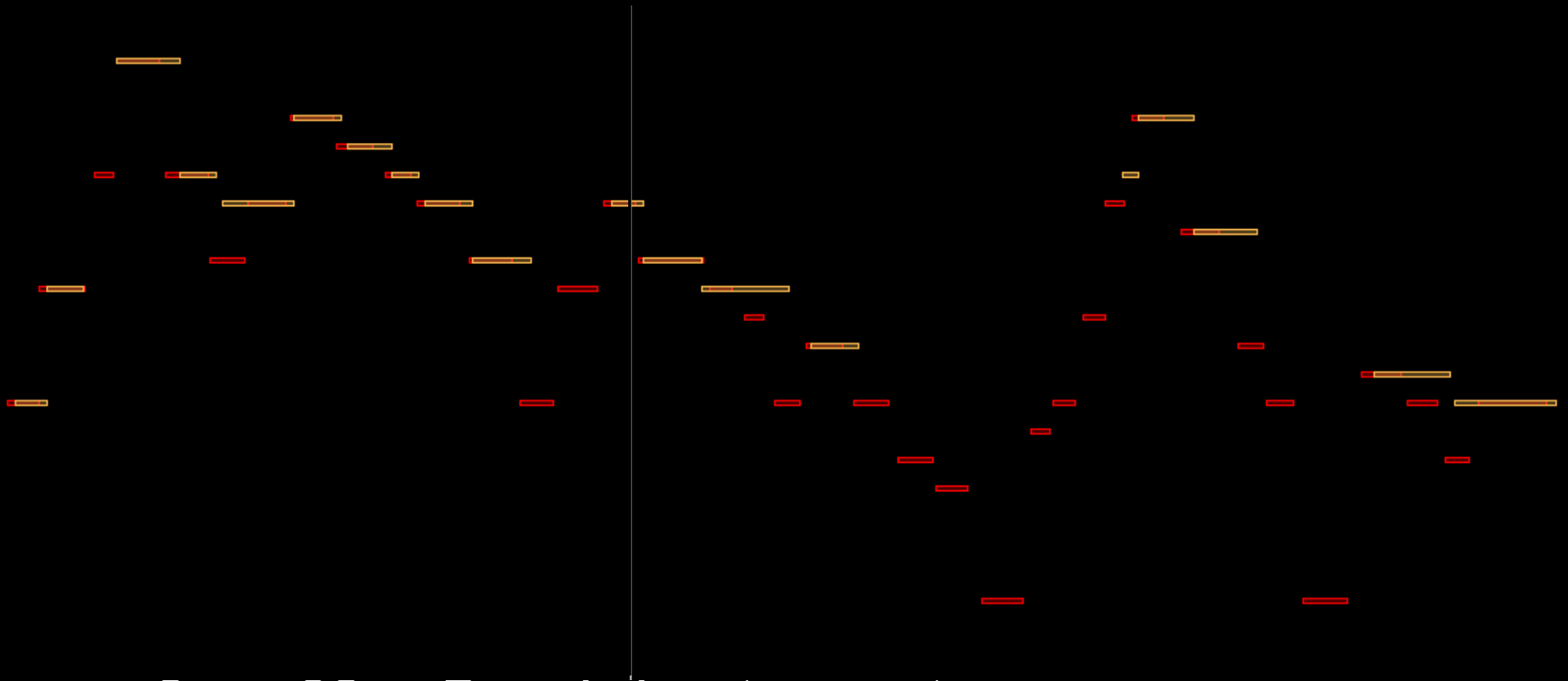
- Viele Transkriptionen verfügbar (z.B. im Netz)
- Oft nicht die, die man will.
- Unterschiedliche Qualität.
- Ansatz: OMR mit SmartScore.
- SmartScore funktioniert relativ gut, aber
  - hoher Nachbearbeitungsbedarf,
  - keine echten Timinginformationen, alle Events müssen von Hand verschoben werden.



# TRANSKRIPTION: SONGS2SEE

- Extraktion mit Songs2See (Spin-off Fraunhofer IDT Ilmenau) ebenfalls recht gut.
- Probleme:
  - Maximal ca. 80 % der Töne erkannt.
  - Variable Ergebnisse je nach Audioqualität.
  - Hoher Nachbearbeitungsbedarf.
  - Onsets oft nicht präzise genug.
  - Oktavfehler.

# TRANSKRIPTION: SONGS2SEE



Songs2See Extraktion (orange)

Finale Transkription (rot)

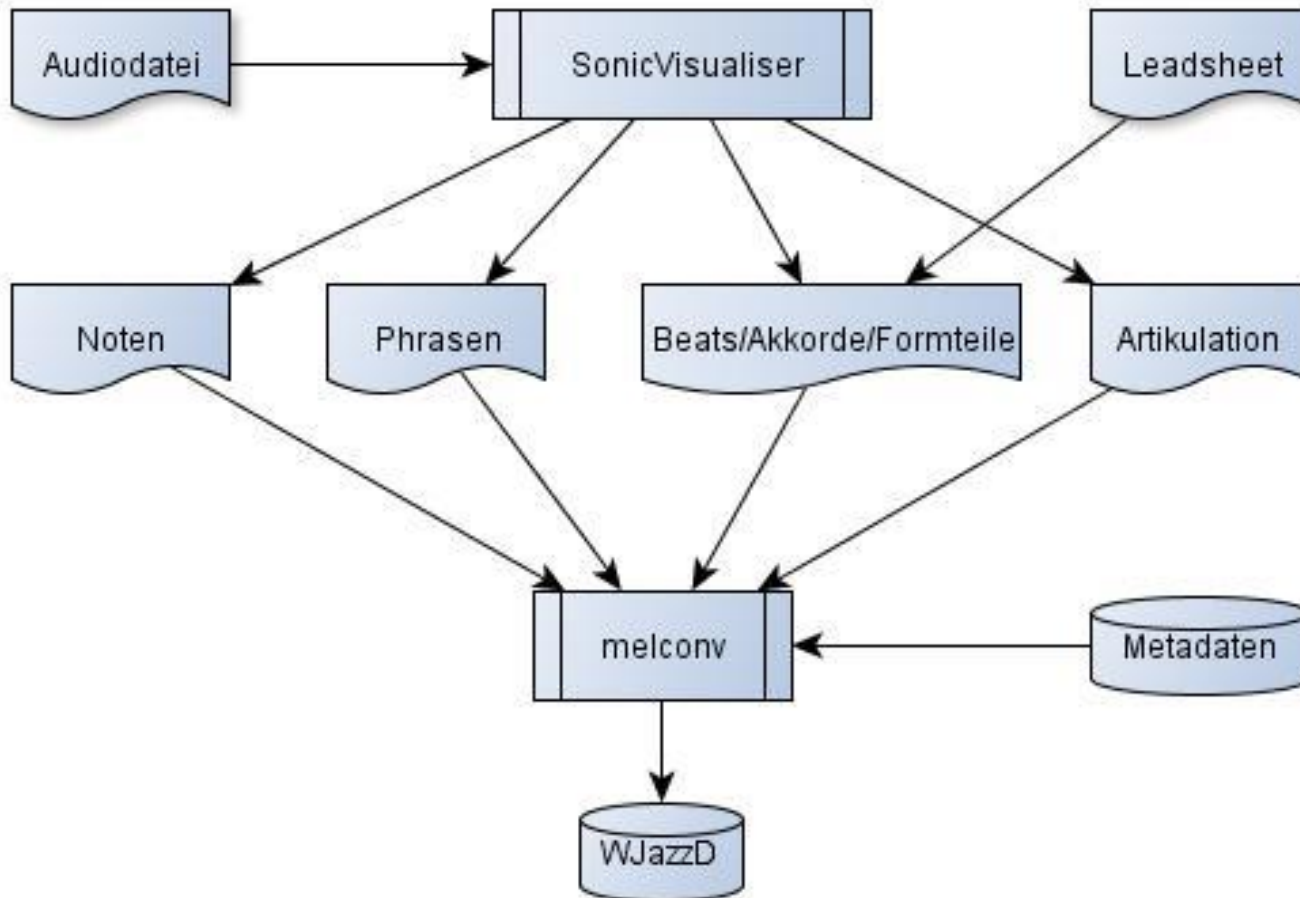
(Sonny Rollins, Blue Seven)

# TRANSKRIPTION: PRAXIS

## Weitere Zugänge:

- Kodierung von Transkriptionen nach EsAC:  
Aufwändig, dann dieselben Probleme wie OMR.
- Einspielen von Transkriptionen per MIDI-Keyboard:  
Aufwändig, fehleranfällig.
- Direkte Annotation im Sonic Visualiser, visuell am  
Spektrogramm mit auditiver Kontrolle:  
**Schnellste und bequemste Technik, die sich  
mittlerweile durchgesetzt hat.**

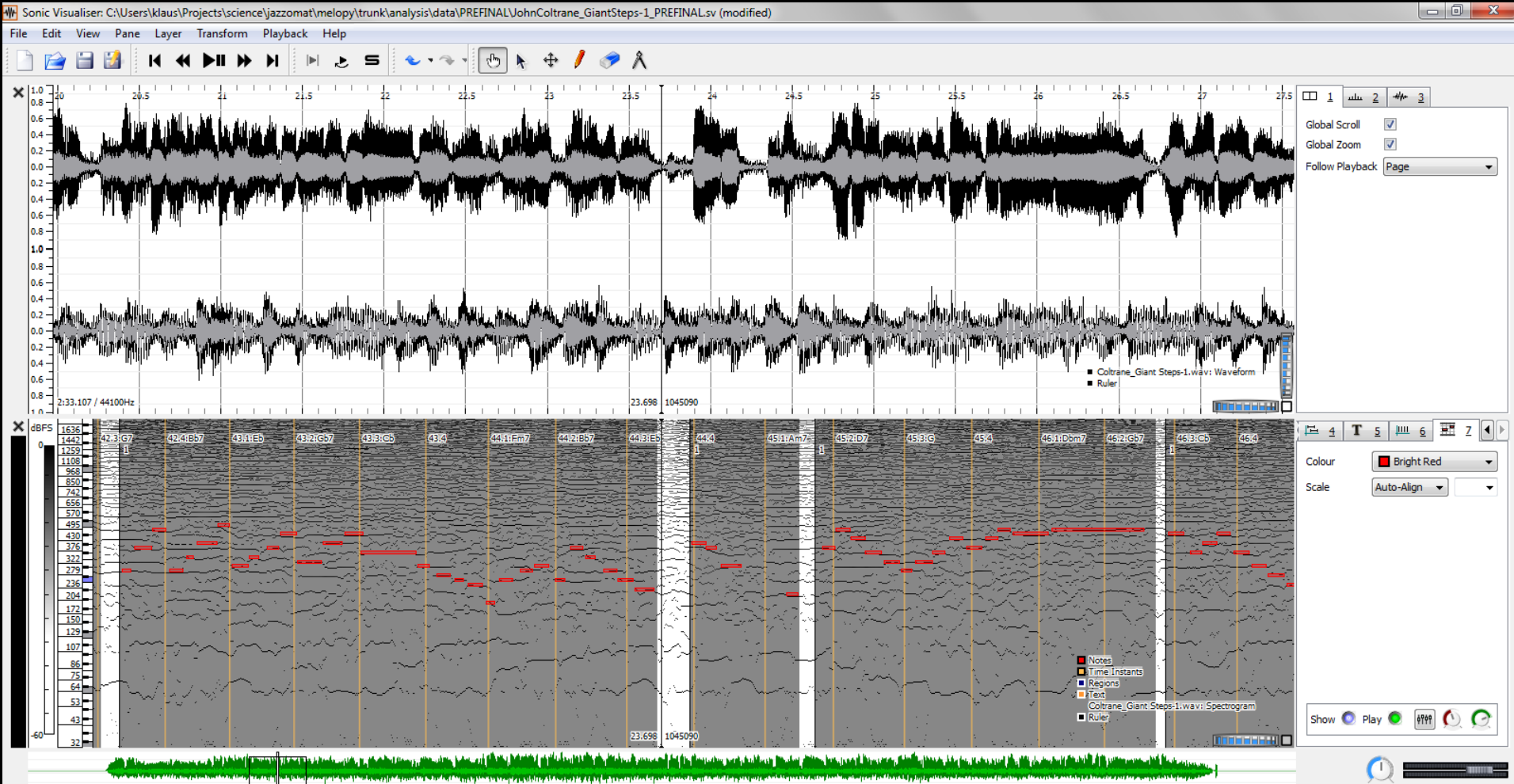
# TRANSKRIPTION: PRAXIS



# EXKURS: SONIC VISUALISER

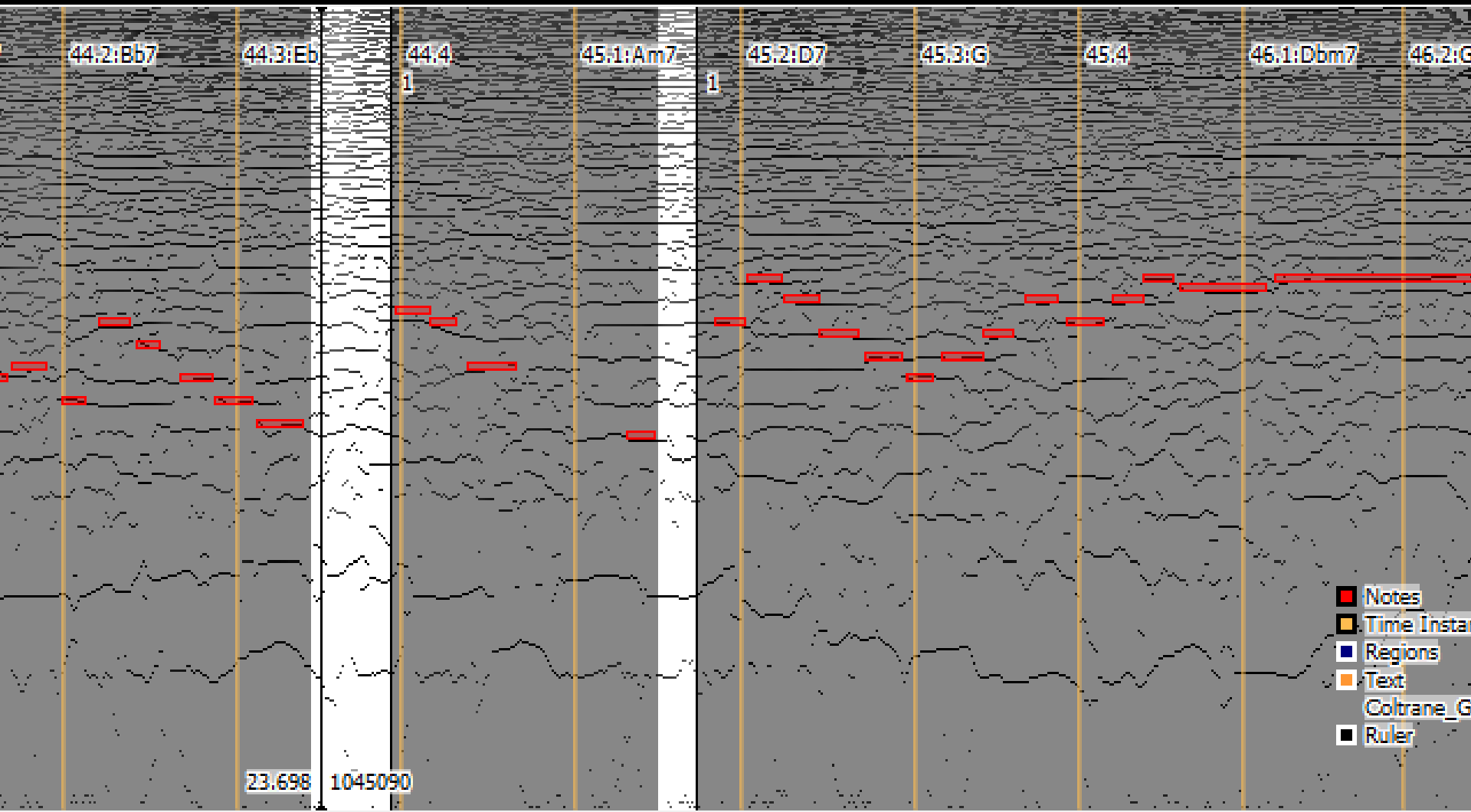
- Frei erhältliches Programm zur Annotation und Visualisierung von Audiodateien
- “The aim of Sonic Visualiser is to be the first program you reach for when want to study a musical recording rather than simply listen to it”
- Layer-Konzept, editierbar, im/exportierbar .
- Integration von MIR-Algorithmen (VAMP Plugins).
- Entwickelt vom Centre for Digital Music, Queen Mary, University of London (Chris Cannam u.a.).
- Relativ ausgereift (V2.13).
- URL: <http://www.sonicvisualiser.org/>

# EXKURS: SONIC VISUALISER



Adjust Colour property of Notes

# EXKURS: SONIC VISUALISER



# METADATEN

- Metadaten in Exceltabelle gesammelt.
- Enthält u.a.:
  - Soloinformationen (Solist, Titel, Stil, Rhythm Feel, Tempoklasse etc.),
  - Kompositionsinformation (Komponist, Form, Tonalitätstyp etc.),
  - Aufnahmeinformationen (Albumtitel, LC Nr, Track Nr., Sidemen, Jahr, MusicBrainzID etc.),
  - Transkriptionsinformationen (wer , wie, wann).



# HARMONISCHER KONTEXT

- Harmonischer Kontext
  - Als Teil der Komposition.
  - Als Leitfaden im Kopf der Solisten.
  - Realisierung durch Harmonieinstrumente.
  - Freiheiten bei der Ausgestaltung (Alterationen, Substitutionen, Wechselakkorde).
- Transkription der klingenden Harmonien schwierig und zeitaufwendig.
- Lösung: Akkordfolgen aus Leadsheet.
- (Ebenso Formteile.)

(MED. W)

## ORNITHOLOGY

- CHARLIE PARKER

G maj7    G-7 C7    G-7 C7  
 F maj7    F-7    Bb7  
 Eb7    D7    2. G-    C-7b5 D7  
 B-7    E7    A-7    D7  
 2. G    A-7 D7    G/B Bb7    A-7 Ab7  
 G maj7    (A-7 D7)  
 FINE

"THE COMPREHENSIVE CHARLIE PARKER"

"LIVE PERFORMANCES-VOL. I"

# METRISCHE ANNOTATION

- Grundlage ist ein kontinuierlicher Beattrack.
- Metrum: Periode des Beats (evtl. Beatproportionen für ungerade Metren).
- Flexibles Gitter: Jedes IBI besitzt eigenes Gitter (optional: nicht-isochrone Tatum).
- Jede Tatumposition kann Subtatum haben (bisher noch nicht eingesetzt).

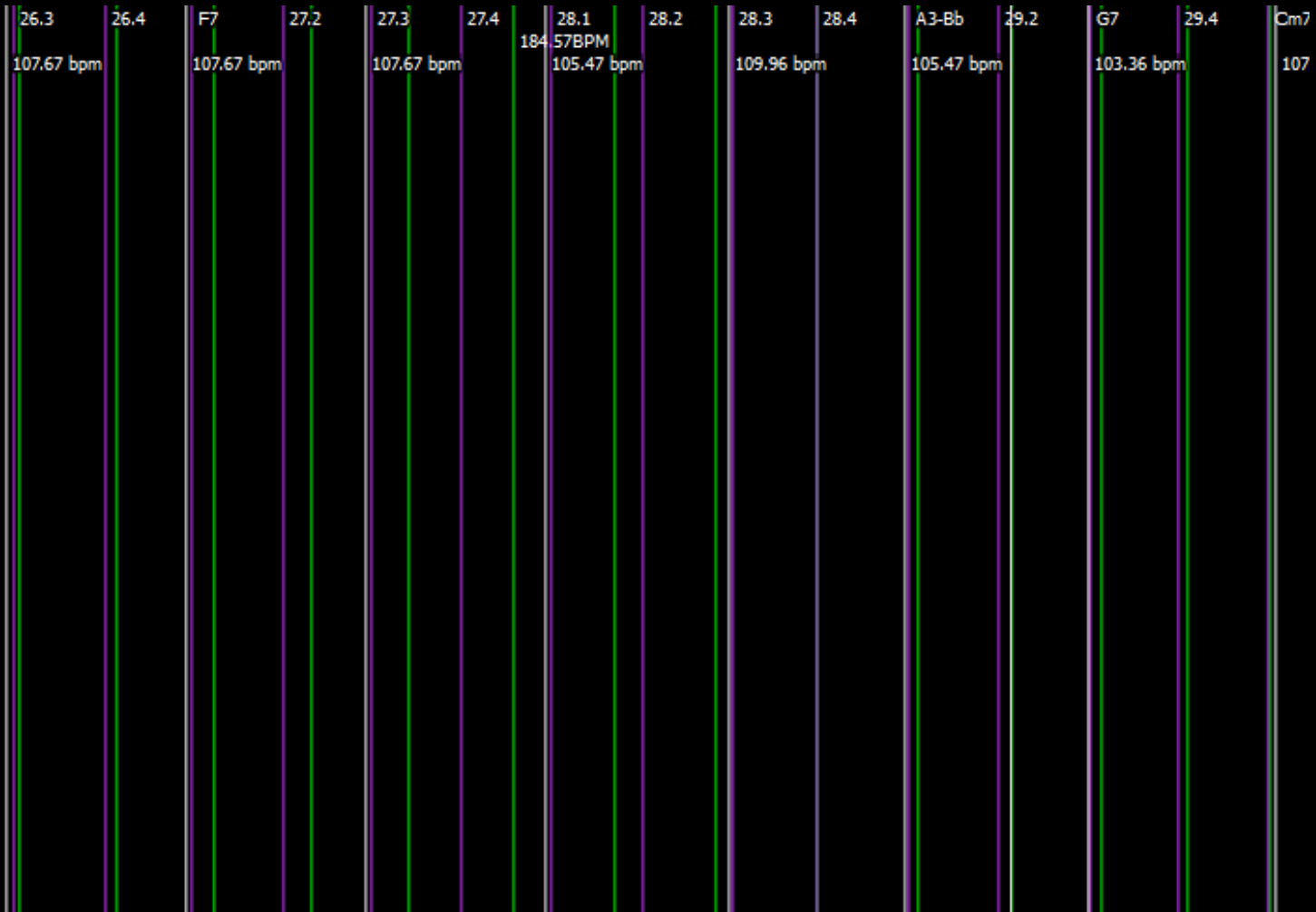
# METRISCHE ANNOTATION

- Vorgehen:
  - Beattrack beschaffen
  - Annotation von Metrum und Phasenlage
  - Bestimmung der Tatumpositionen per Algorithmus (FlexQ)
- Wie kommt man da ran?

# BEATTRACKING

- Verschiedene Beattracker ausprobiert (SV Plugins).
- Keiner war hinreichend gut: Hoher Nachbearbeitungsbedarf.
- Typische Probleme:
  - Präzision,
  - Halftime bei hohem Tempo,
  - Phasenverschiebungen.
- Lösung: Manuelles Tappen

# BEATTRACKING



Purple: Manuell. Weiß: QMUL Beatracker, Grün: IBT-INESC Beatracker

# BEATTRACKING

- Probleme des Beattappen:
  - Asynchronie,
  - Computer-Latenzen,
  - Fokussierung,
  - Präzision (insb. bei hohem Tempo).
- Wo ist der Beat eigentlich? Wessen Beat eigentlich?
  - Laid back/Ahead-Spiel,
  - interne Asynchronien,
  - PATs.

# PHRASEN

- Phrasen sind wichtige perzeptuelle und produktive musikalische Einheiten, interessant für Analyse.
- Ansatz: Transkripteur annotiert Phrasen.
- Probleme:
  - Variabilität in der Phrasenwahrnehmung
  - Nur ein Transkripteur, wenig reliabel
- Geplant:
  - Validierung durch Laborexperimente,
  - Segmentierungsalgorithmus für Jazzsoli.



# ARTIKULATION

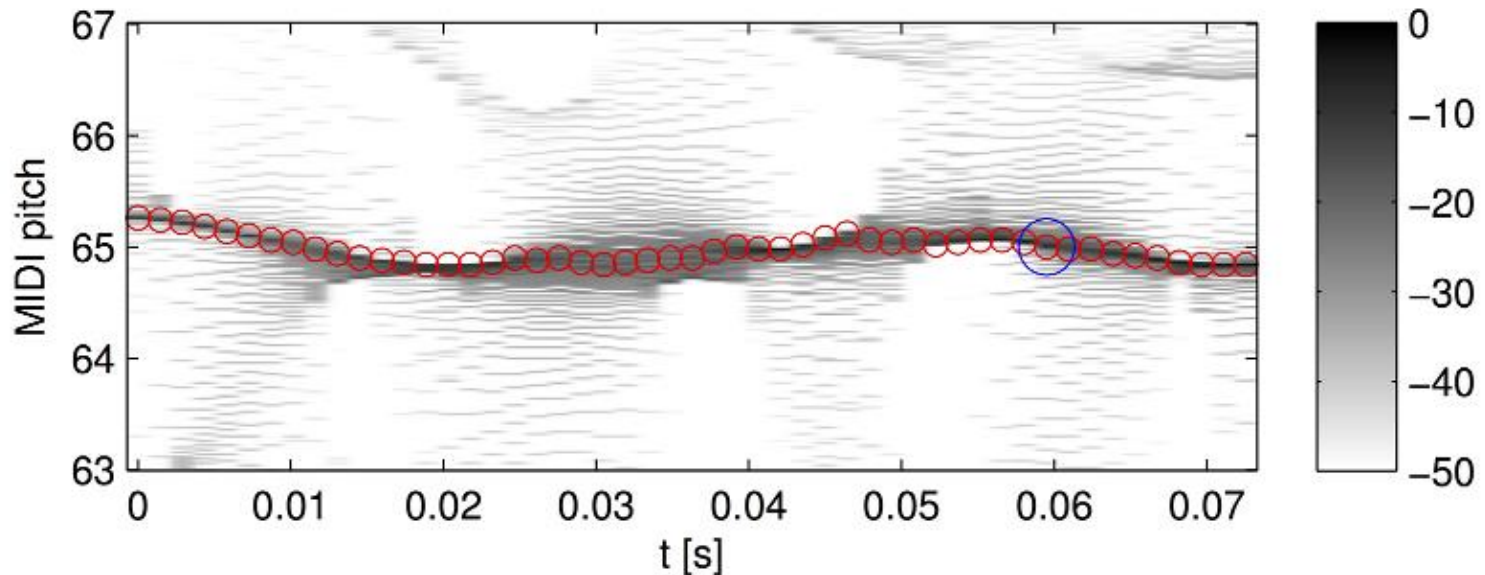
- Artikulation wichtiger expressiver und performativer Parameter.
- Hier als Modulationen des Tonhöhenverlaufs (nicht als Staccato, Legato etc. → Dauern/OOIs).
- Manuelle Annotation durch die Transkripteure (teilweise).
- Klassen: Vibrato, Bend, Fall-off, Slide.
- Nicht berücksichtigt: Multiphonics, Flageolets, Ghostnotes, False-Fingerings.

# LAUTHEIT & ARTIKULATIONSMESSUNG

- Laufendes Projekt
- **Idee:** Transkriptionen als Leitfaden für  $f_0$ -Tracking und Intensitätsmessung
- Kennzahlen der  $f_0$ -Kurven: Vibratofrequenz, Vibratohub,  $f_0$ -Gradient etc.
- Erlaubt auch Intonationsmessung.
- Nachträgliche automatisierte Annotation von Intensitäten.

# ARTIKULATIONSMESSUNG

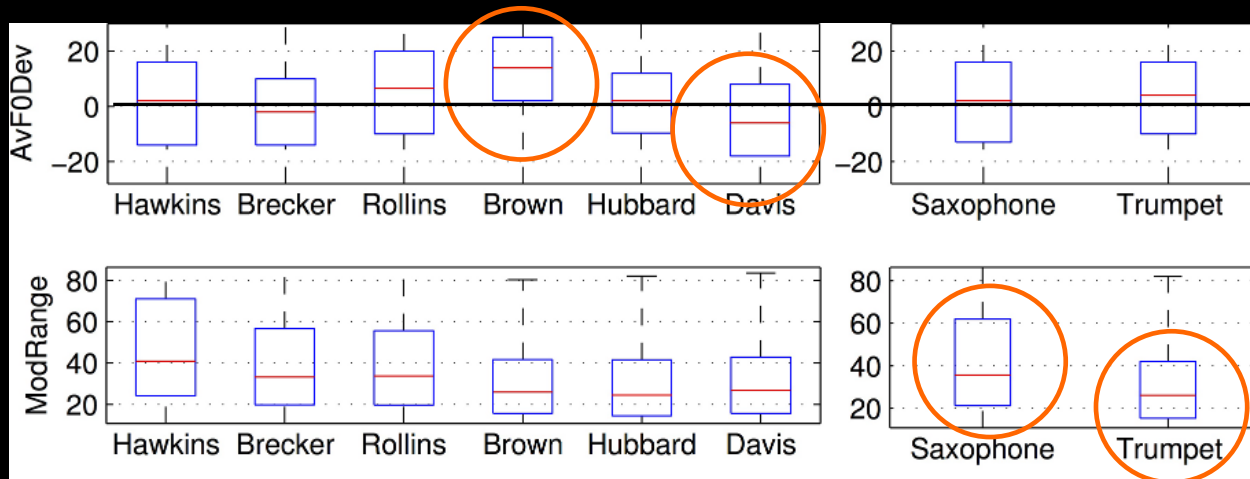
- Beispiel:  $f_0$ -Track, Vibrato



Stompin' At The Savoy, Coleman Hawkins (ts)

# ARTIKULATIONSMESSUNG

- Intonation: Z.B. Clifford Brown: sharp, Miles Davis: flat.
- Saxophonisten mehr Vibratohub als Trompeter,
- Vibratohub nimmt mit Tonhöhe ab ( $r=.28^{***}$ )



# ZUSAMMENFASSUNG

- Die Herstellung von maschinenverwertbaren Transkriptionen erfordert zahlreiche theoretische und pragmatische Entscheidungen.
- MIR-Algorithmen zur Transkription recht gut, aber für praktische Anwendung dieser Art (noch) nicht gut genug.
- Sonic Visualiser zur Annotation sehr geeignet.
- Semi-automatische Zugänge oft das Mittel der Wahl.
- Transkriptionen erlauben MIR-Algorithmen zu informieren und zu verbessern.

# AUSBLICK: JAZZKARAOKE MIT CHARLIE

Melody



Charlie Parker  
Ornithology 1946





VIELEN DANK!