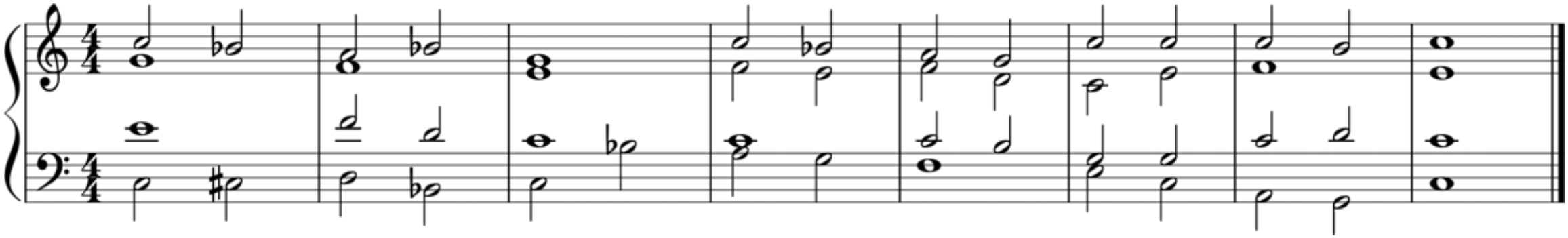
On Local Keys, Modulations, and Tonicizations:

A Dataset and Methodology for Evaluating Changes of Key

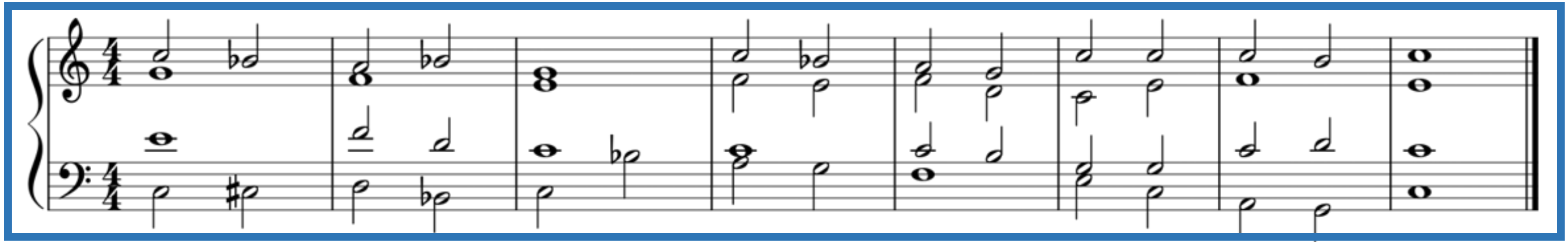
Néstor Nápoles López, Laurent Feisthauer,
Florence Levé, Ichiro Fujinaga

7th International Conference on Digital Libraries for Musicology, 16 October 2020

Key estimation

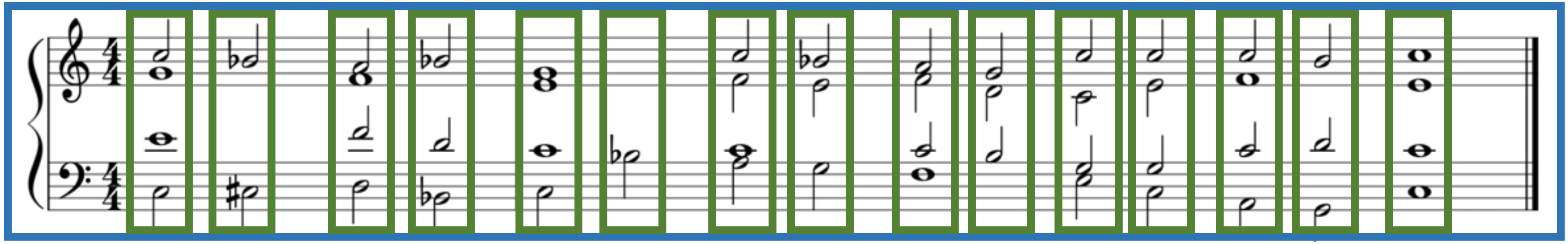


Key estimation

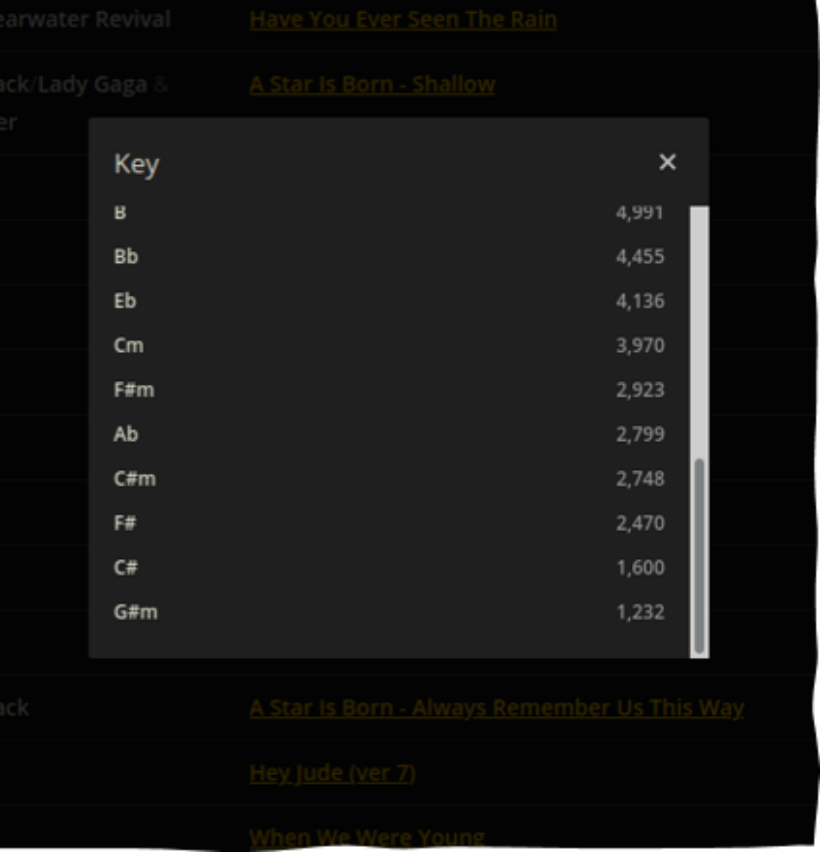


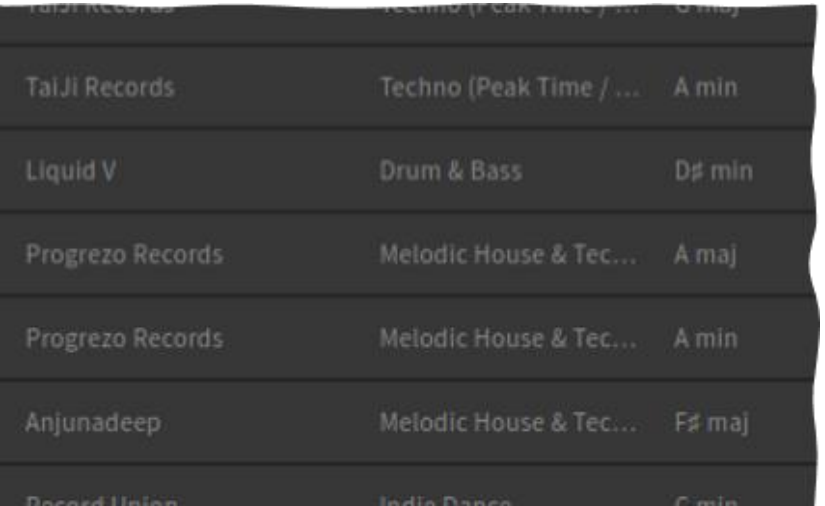
- **Global key estimation**

Key estimation



- **Global key estimation**
- **Local key estimation**





Global key and their applications

Search parameters in a digital library :

- Ultimate guitar
- Beat port

Local keys and their applications

Used mostly as mid-level MIR features for other tasks:

- Roman numeral analysis (Temperley, 2004; Micchi et al., 2020)
- MIDI pitch spelling (Teodoru et Raphael, 2007)
- Music generation (Hadjeres et al., 2017)

Potentially usable in digital music libraries too:

- Search for pieces with similar modulations or tonicizations patterns

Local keys, modulations, and tonicizations

Outline the necessity for studying local keys more in depth

Particularly, the relationship to concepts like modulation and tonicization

Change of key

- Modulation
- Tonicization
- Local keys

Local keys, modulations, and tonicizations

Outline the necessity for studying local keys more in depth

Particularly, the relationship to concepts like modulation and tonicization

Change of key – tonal music theory

- Modulation
- Tonicization
- Local keys????

Local keys, modulations, and tonicizations

Outline the necessity for studying local keys more in depth

Particularly, the relationship to concepts like modulation and tonicization

Change of key – music information retrieval

- Modulation????
- Tonicization????
- Local keys

Local keys, modulations, and tonicizations

Outline the necessity for studying local keys more in depth

Particularly, the relationship to concepts like modulation and tonicization

Change of key

- Modulation
- Tonicization
- Local keys

Investigate!

This paper

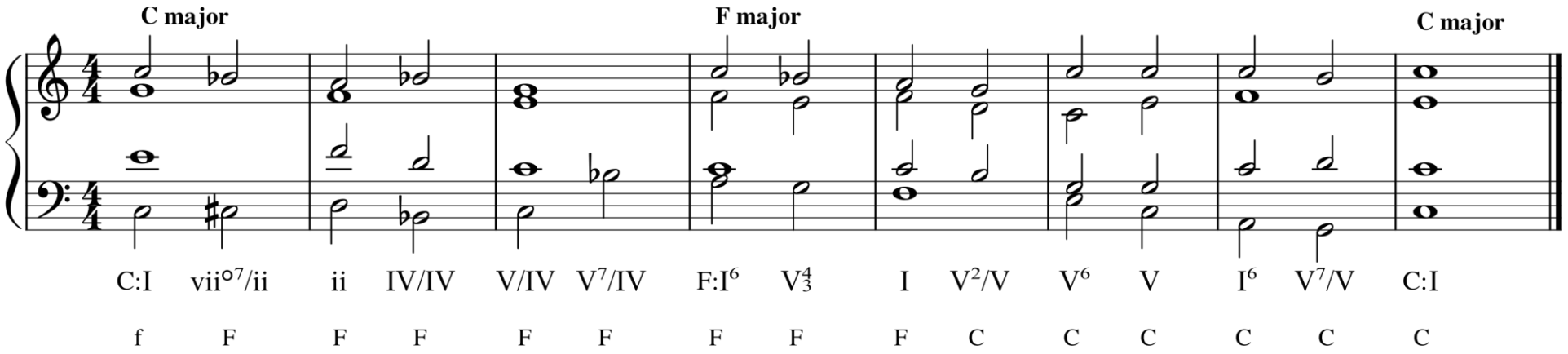
Method for encoding modulation/tonicization ground truth

Dataset with 201 excerpts (five tonal music theory textbooks)

Comparing local-key predictions with modulation and tonicization ground truth annotations

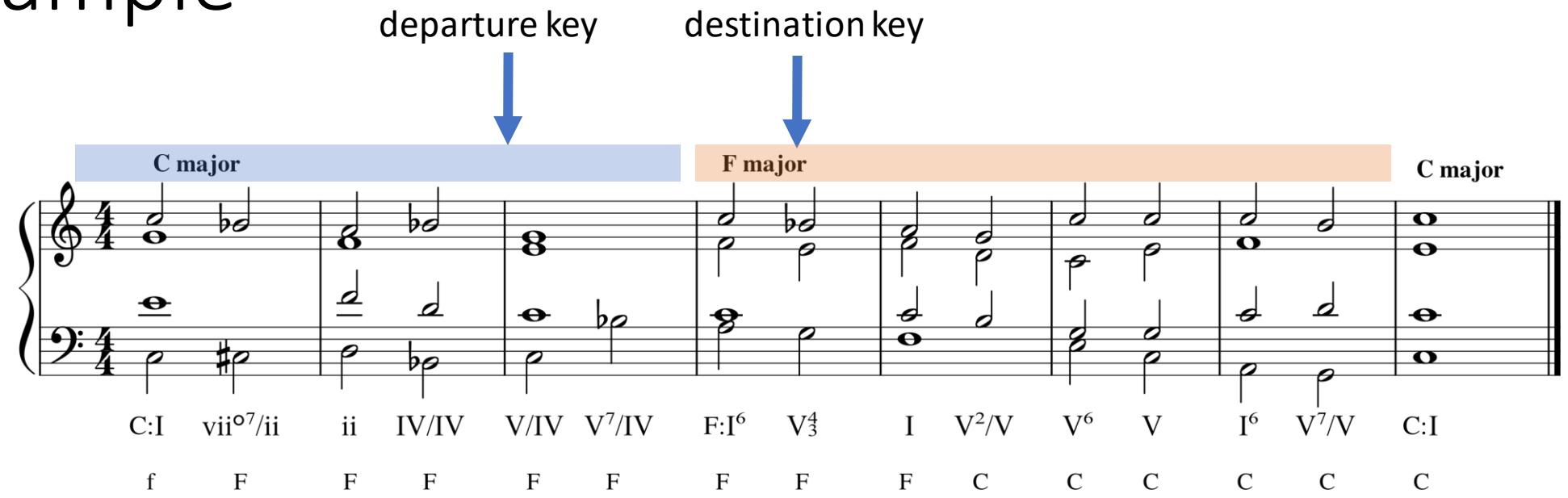
Using the data and methodology, evaluate three local-key estimation models from the literature

Example



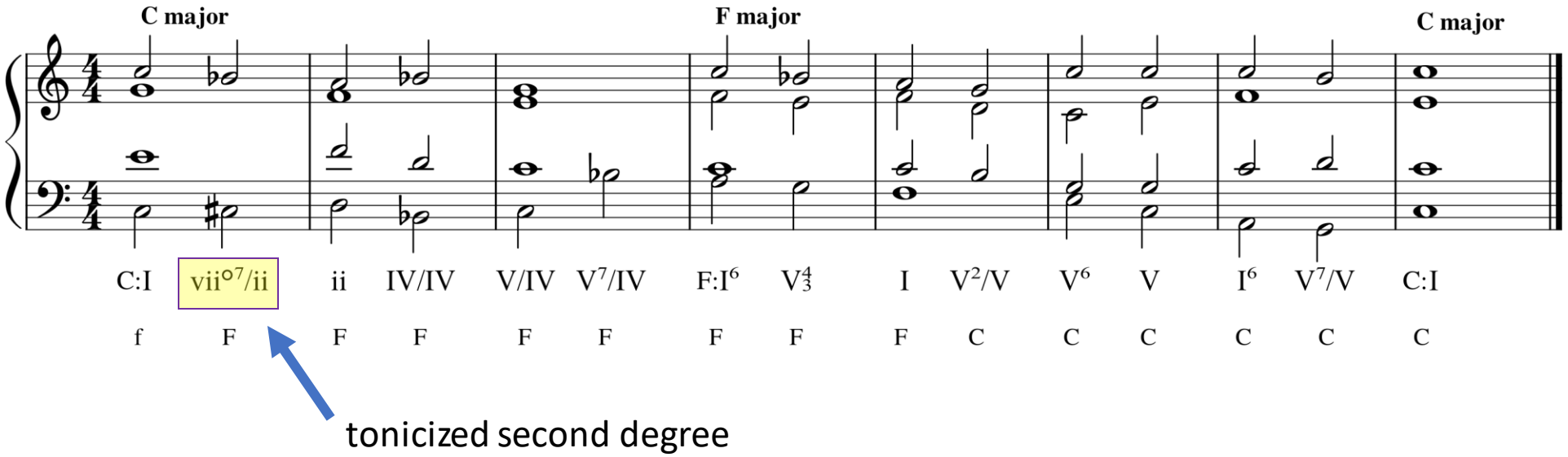
1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

Example



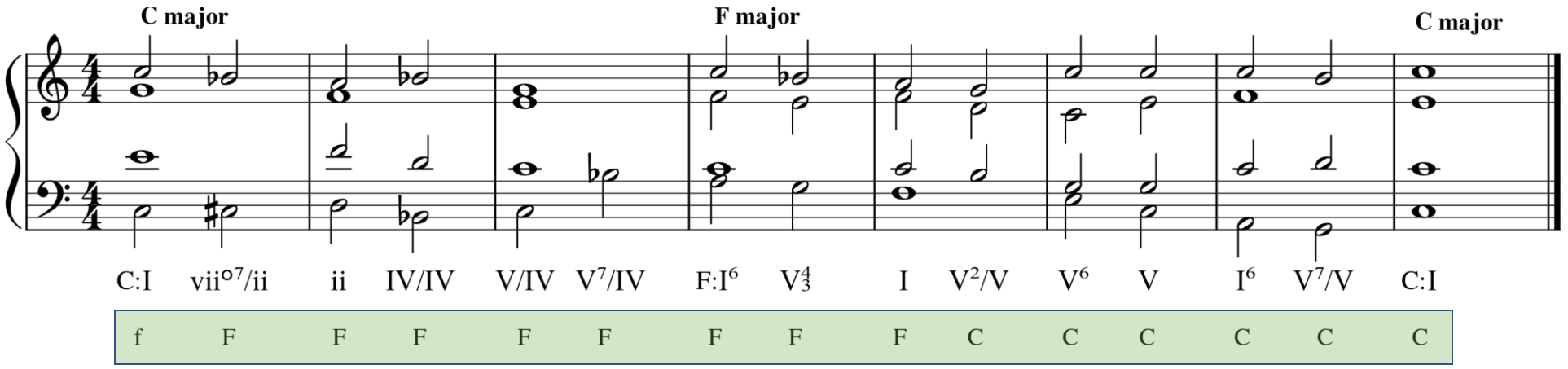
1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

Example



1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

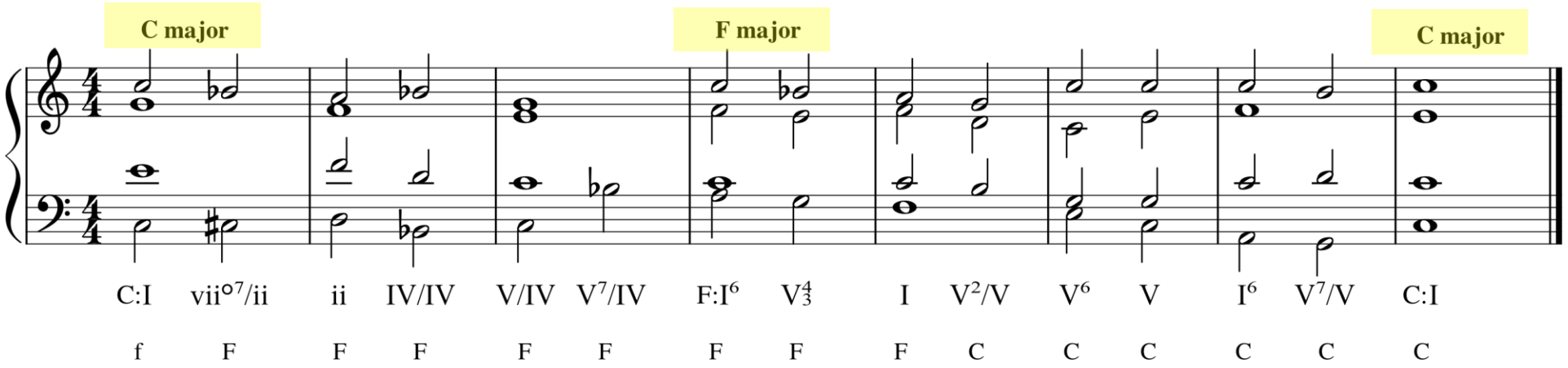
Example



local key predictions at every onset

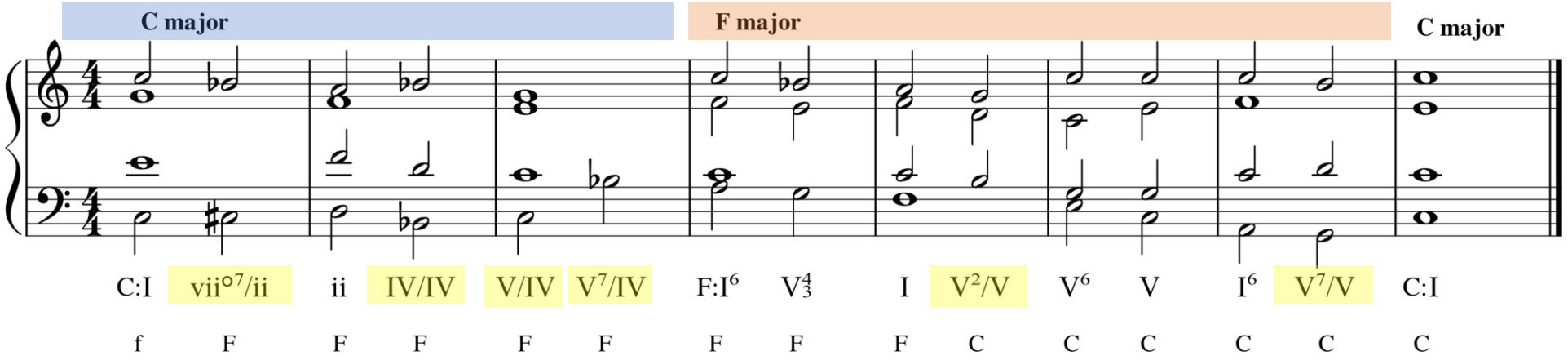
1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

Example



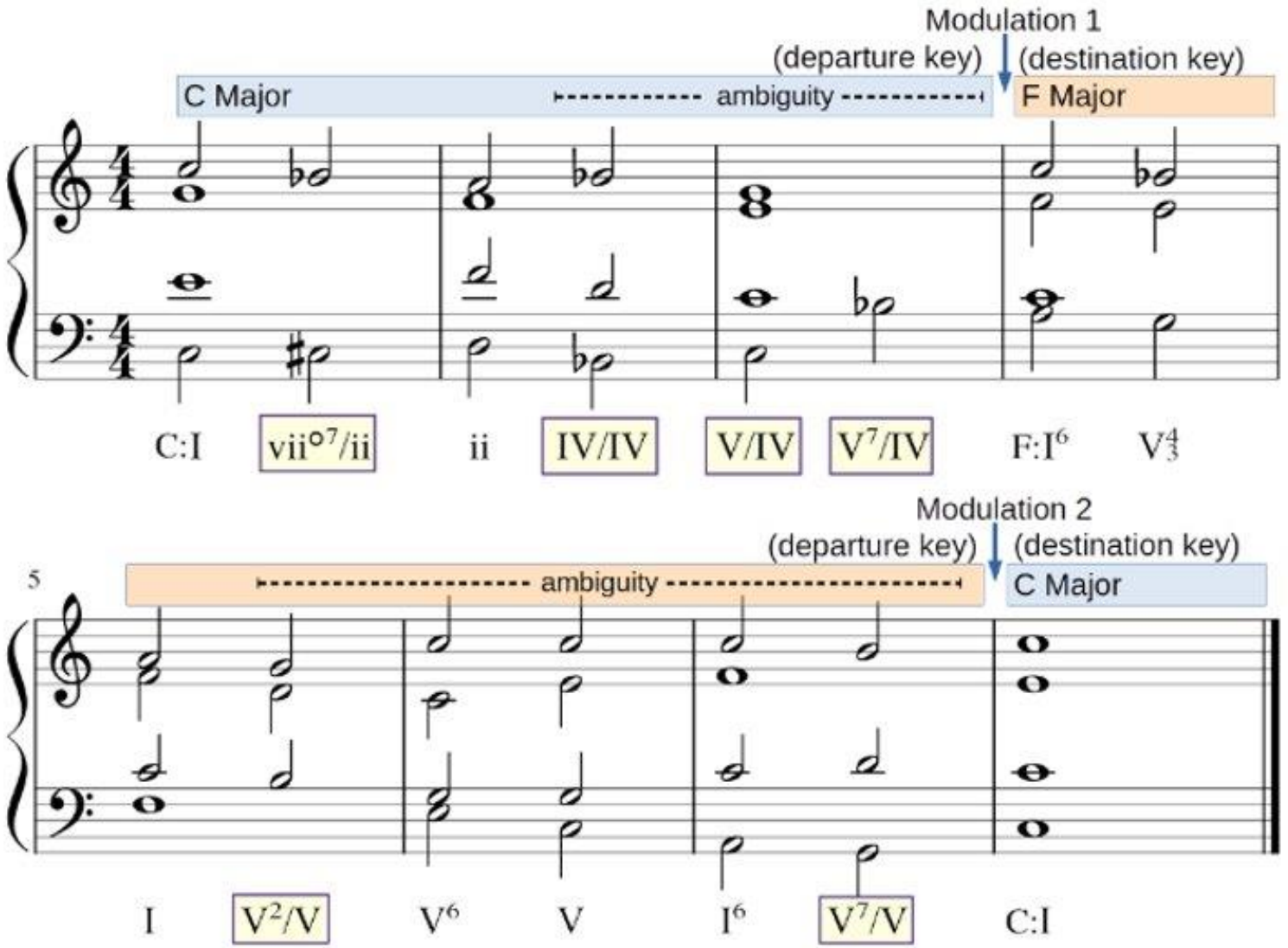
1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

Example



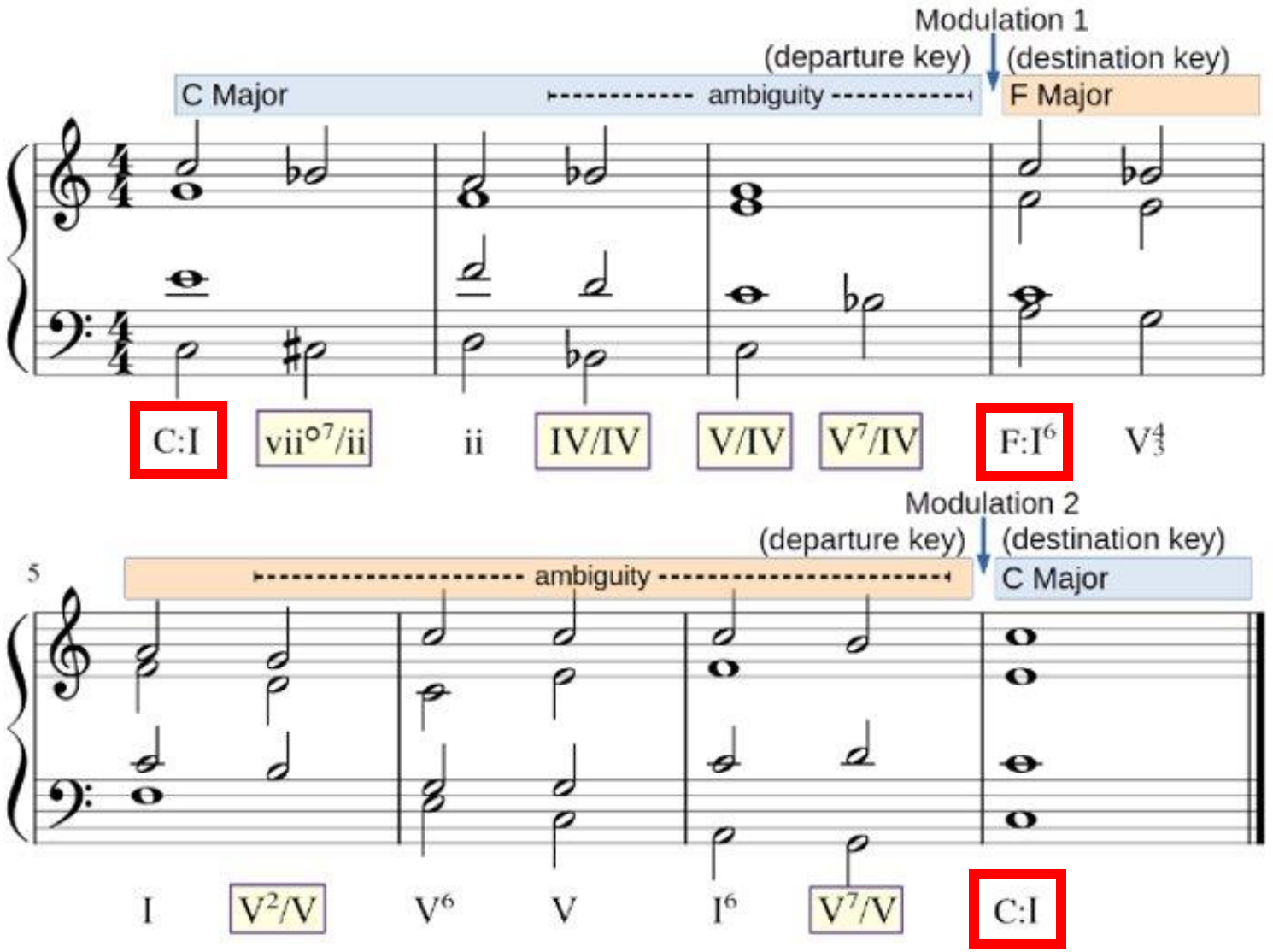
1. Modulations are **always** taken from the textbooks
2. Tonicizations are inferred from roman numerals
3. Roman numeral analysis is **not always** provided in the textbooks

Example



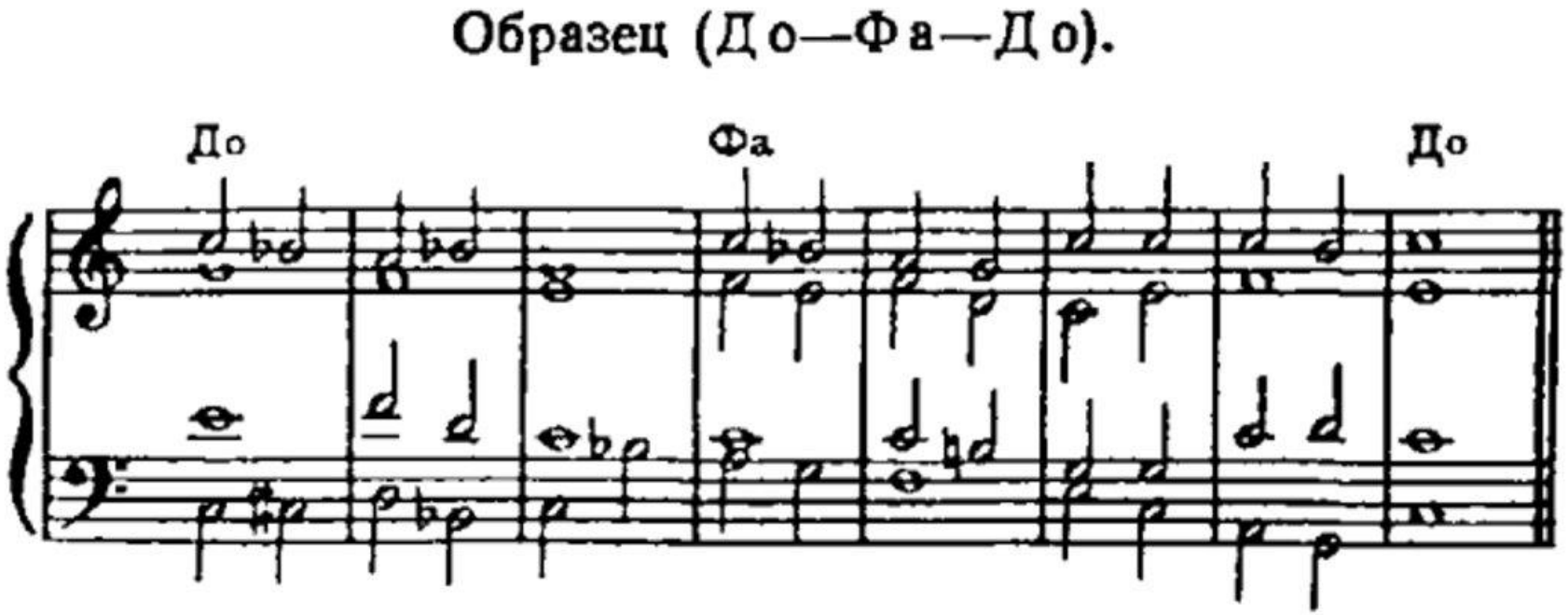
Example 3-17b in Rimsky-Korsakov's *Practical Manual of Harmony*

Example



Example 3-17b in Rimsky-Korsakov's *Practical Manual of Harmony*

Original edition by Rimsky-Korsakov

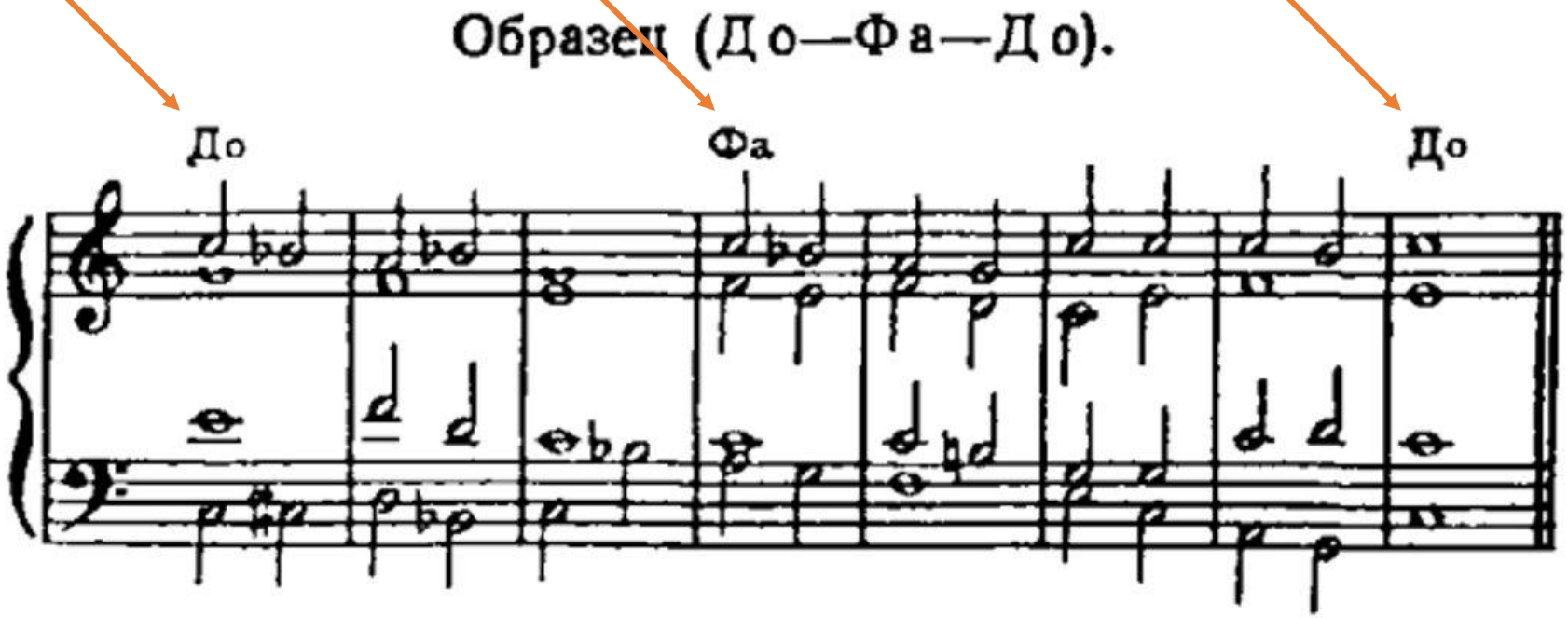


Original edition by Rimsky-Korsakov

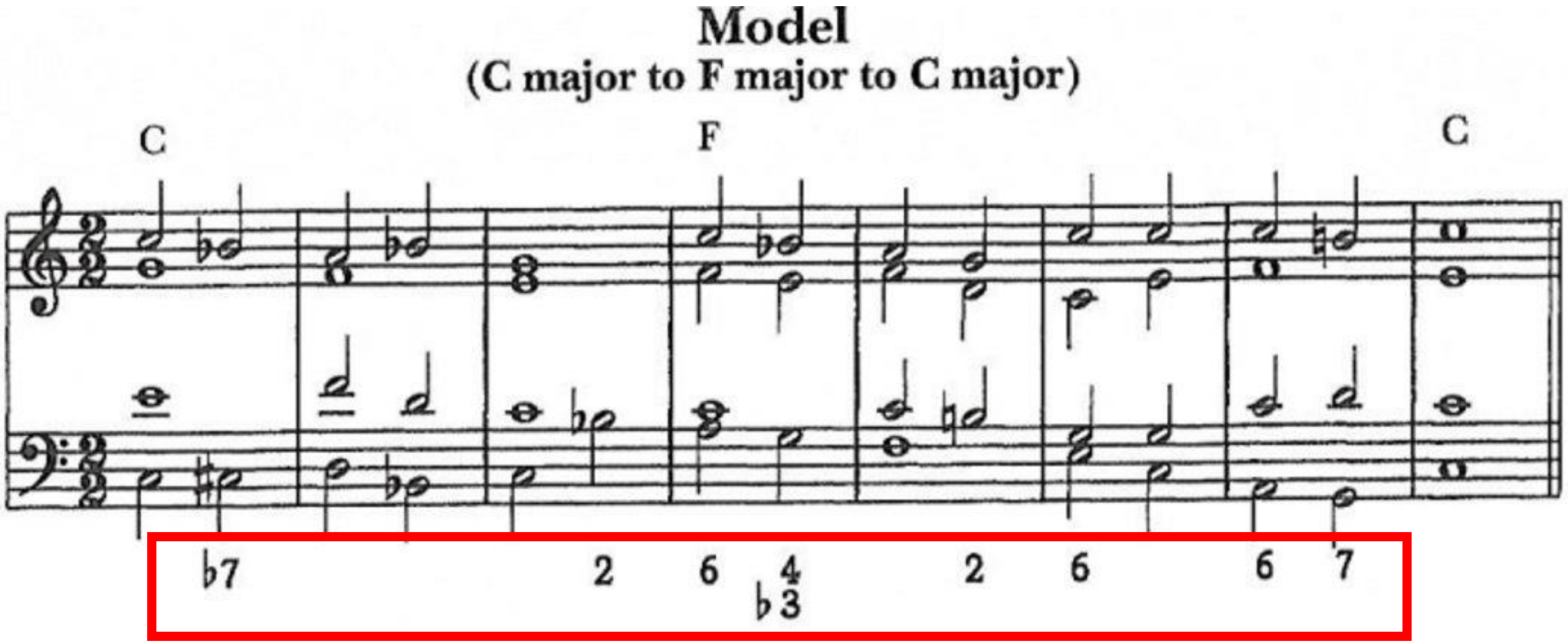
C Major

F Major

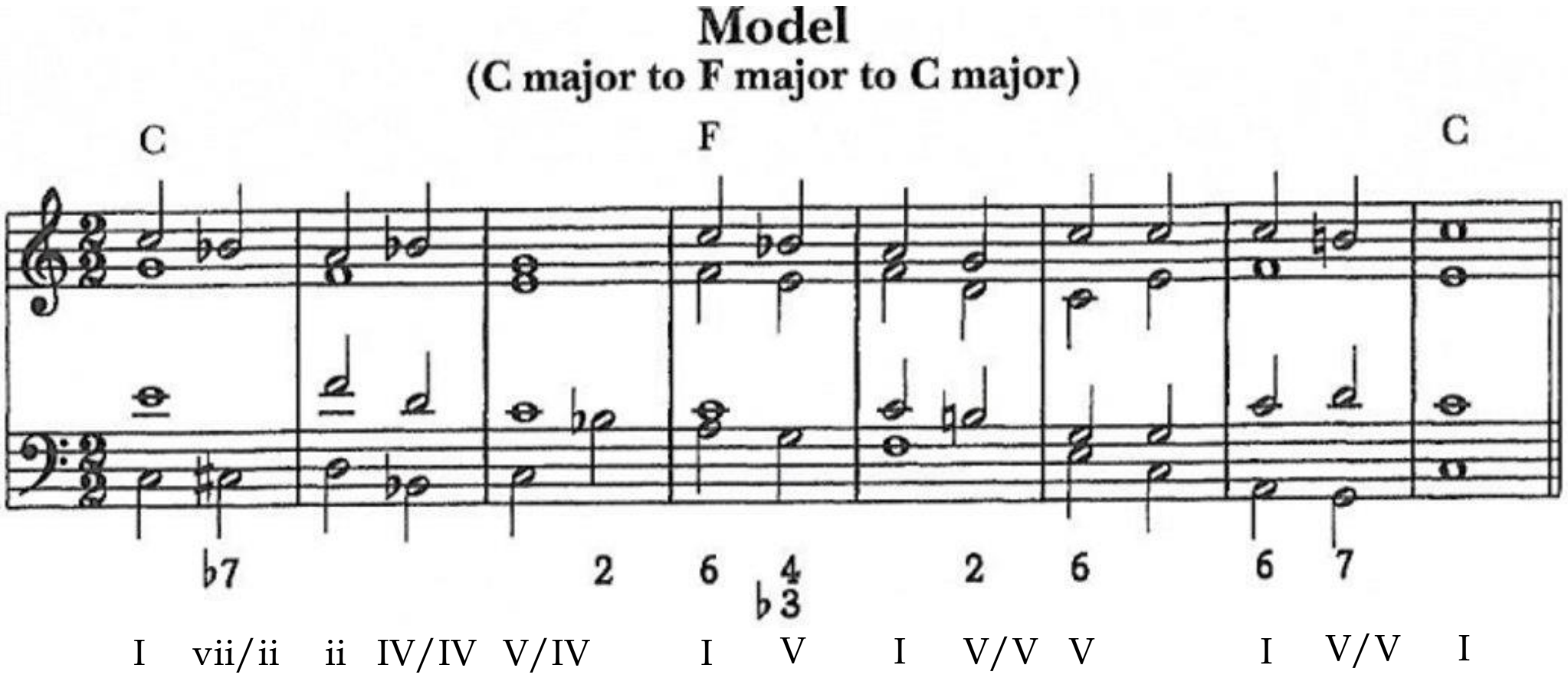
C Major



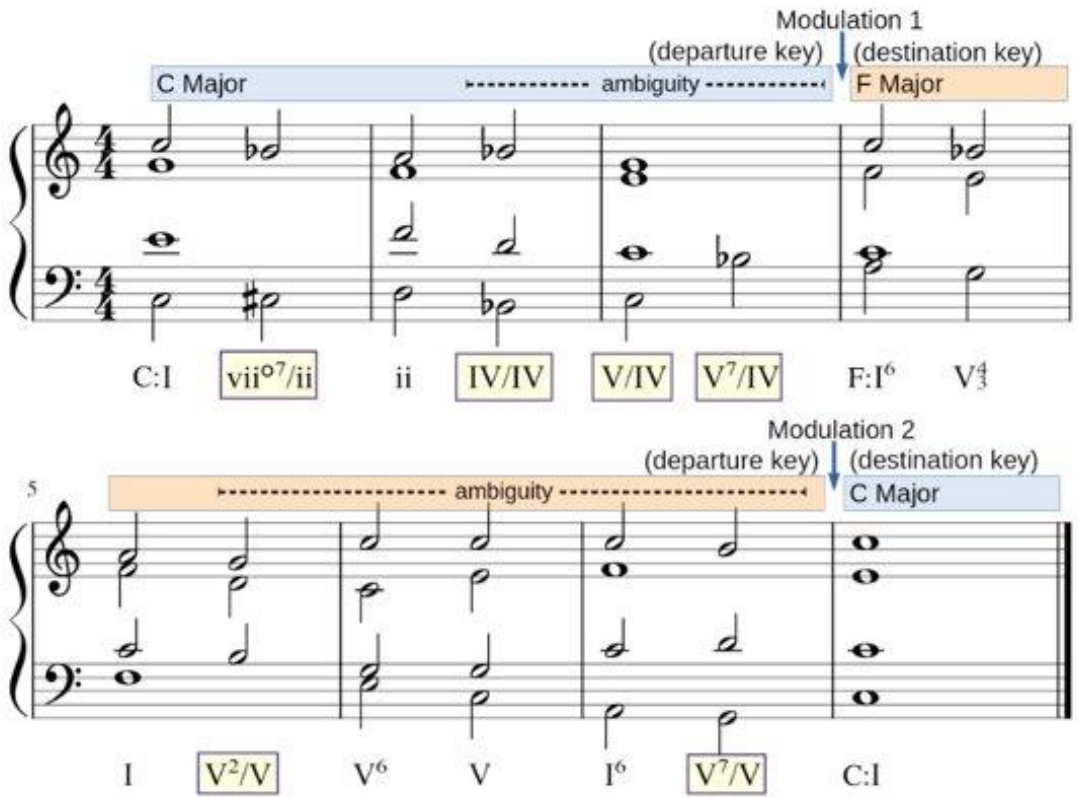
The one we used



The one we used



Processing annotations



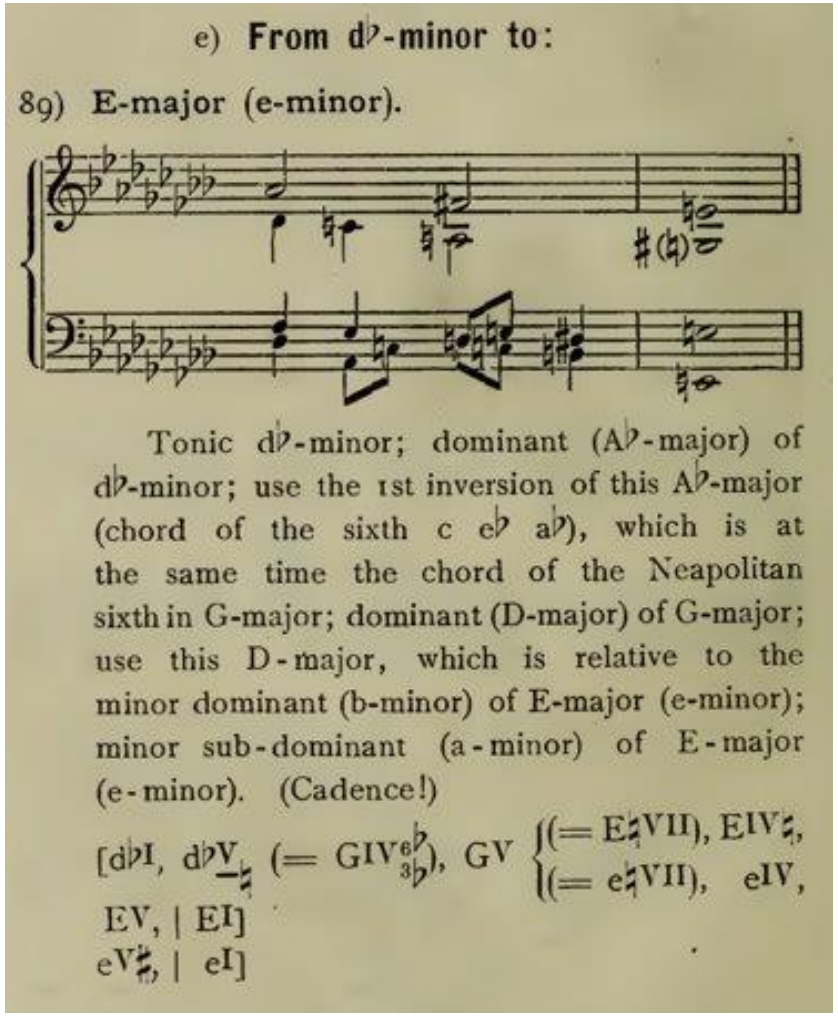
Position	Annotation	Modulation	Tonicization
0	C=>:I	C major	C major
2	viio7/ii	C major	D minor
4	ii	C major	C major
6	IV/IV	C major	F major
8	V/IV	C major	F major
10	V7/IV	C major	F major
12	F=>:I6	F major	F major
14	V43	F major	F major
16	I	F major	F major
18	V2/V	F major	C major
20	V6	F major	F major
22	V	F major	F major
24	I6	F major	F major
26	V7/V	F major	C major
28	C=>:I	C major	C major

Dataset

- (ASC) Aldwell, Cadwallader and Schachter
  Harmony and Voice Leading
- (KP) Kostka and Payne
  Tonal Harmony
- (Reg) Reger
  On the Theory of Modulation
- (Rim) Rimsky-Korsakov
  Practical Manual of Harmony
- (Tch) Tchaikovsky
  Guide to the practical study of Harmony

Sample	Files	Modulations	Tonicizations	Labels
ASC	7	8	7	185
KP	15	21	11	554
Reg	117	220	40	768
Rim	37	44	107	257
Tch	25	60	38	238
Total	201	555	203	2002

Some of the textbooks



e) From d♭-minor to:

89) E-major (e-minor).

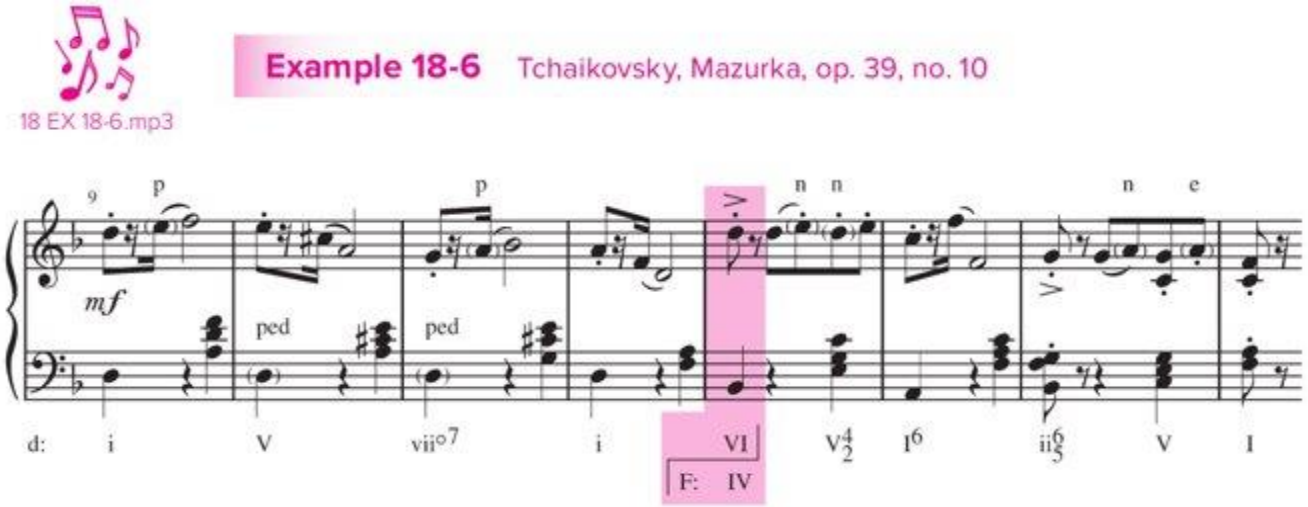
Tonic d♭-minor; dominant (A♭-major) of d♭-minor; use the 1st inversion of this A♭-major (chord of the sixth c e♭ a♭), which is at the same time the chord of the Neapolitan sixth in G-major; dominant (D-major) of G-major; use this D-major, which is relative to the minor dominant (b-minor) of E-major (e-minor); minor sub-dominant (a-minor) of E-major (e-minor). (Cadence!)

[d♭I, d♭V (= GIV$^{6}_{3}$♭), GV {(= E♮VII), EIV♮, (= e♮VII), eIV,

EV, | EI]

eV♯, | eI]

Extract from Reger



Extract from Kostka and Payne

Models

Baseline

- (B1) Random guess
- (B2) Global key

MIR models

- (M1) Nápoles López et al. (2019)
- (M2a-b) Feisthauer et al. (2020)
- (M3) Micchi et al. (2020)

Position	Annotation	Modulation	Tonicization
0	C:=>I	C	C
2	viio7/ii	C	d
4	ii	C	C
6	IV/IV	C	F
8	V/IV	C	F
10	V7/IV	C	F
12	F:=>I6	F	F
14	V43	F	F
16	I	F	F
18	V2/V	F	C
20	V6	F	F
22	V	F	F
24	I6	F	F
26	V7/V	F	C
28	C:=>I	C	C

Models

Baseline

- (B1) Random guess
- (B2) Global key

MIR models

- (M1) Nápoles López et al. (2019)
- (M2a-b) Feisthauer et al. (2020)
- (M3) Micchi et al. (2020)

Position	Annotation	Modulation	Tonicization	B1
0	C:=>I	C	C	Bb
2	viio7/ii	C	d	a
4	ii	C	C	Bb
6	IV/IV	C	F	eb
8	V/IV	C	F	C#
10	V7/IV	C	F	f
12	F:=>I6	F	F	C#
14	V43	F	F	g
16	I	F	F	c
18	V2/V	F	C	C#
20	V6	F	F	C#
22	V	F	F	g
24	I6	F	F	b
26	V7/V	F	C	e
28	C:=>I	C	C	d

Models

Baseline

- (B1) Random guess
- (B2) Global key

MIR models

- (M1) Nápoles López et al. (2019)
- (M2a-b) Feisthauer et al. (2020)
- (M3) Micchi et al. (2020)

Position	Annotation	Modulation	Tonicization	B1	B2
0	C:=>I	C	C	Bb	F
2	viio7/ii	C	d	a	F
4	ii	C	C	Bb	F
6	IV/IV	C	F	eb	F
8	V/IV	C	F	C#	F
10	V7/IV	C	F	f	F
12	F:=>I6	F	F	C#	F
14	V43	F	F	g	F
16	I	F	F	c	F
18	V2/V	F	C	C#	F
20	V6	F	F	C#	F
22	V	F	F	g	F
24	I6	F	F	b	F
26	V7/V	F	C	e	F
28	C:=>I	C	C	d	F

Models

Baseline

- (B1) Random guess
- (B2) Global key

MIR models

- (M1) Nápoles López et al. (2019)
- (M2a-b) Feisthauer et al. (2020)
- (M3) Micchi et al. (2020)

Position	Annot.	Mod.	Ton.	B1	B2	M1	M2 a	M2 b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Accuracy and MIREX

Key Relationship (Reference, Predicted)	Accuracy	MIREX
Same key	1.0	1.0
Dominant / SubDominant	0.0	0.5
Relative Major / Relative Minor	0.0	0.3
Parallel Major / Parallel Minor	0.0	0.2
Other	0.0	0.0

Example :

Ground truth key	Prediction	Accuracy	MIREX
C	F	0.0	0.5

Accuracy and MIREX

Key Relationship (Reference, Predicted)	Accuracy	MIREX
Same key	1.0	1.0
Dominant / SubDominant	0.0	0.5
Relative Major / Relative Minor	0.0	0.3
Parallel Major / Parallel Minor	0.0	0.2
Other	0.0	0.0

Example :

Ground truth key	Prediction	Accuracy	MIREX
C	F	0.0	0.5

Accuracy and MIREX

Key Relationship (Reference, Predicted)	Accuracy	MIREX
Same key	1.0	1.0
Dominant / SubDominant	0.0	0.5
Relative Major / Relative Minor	0.0	0.3
Parallel Major / Parallel Minor	0.0	0.2
Other	0.0	0.0

Example :

Ground truth key	Prediction	Accuracy	MIREX
C	F	0.0	0.5

Comparison of the models and the ground truth

Local keys vs. Modulation

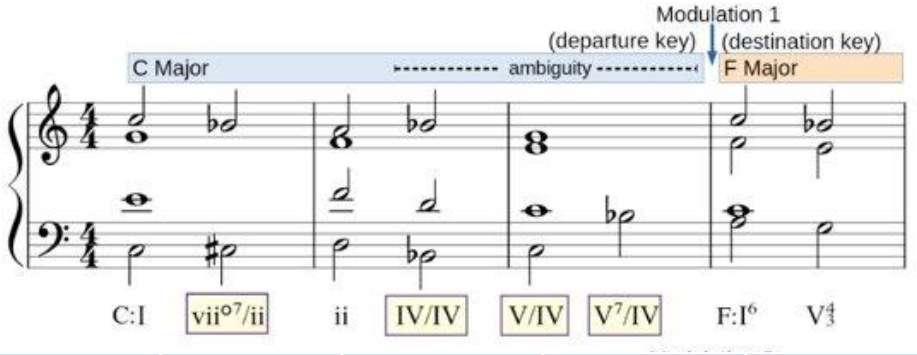
- Accuracy weights
- MIREX weights

Local keys vs. Tonicization

- Accuracy weights
- MIREX weights

What do local keys coincide with?

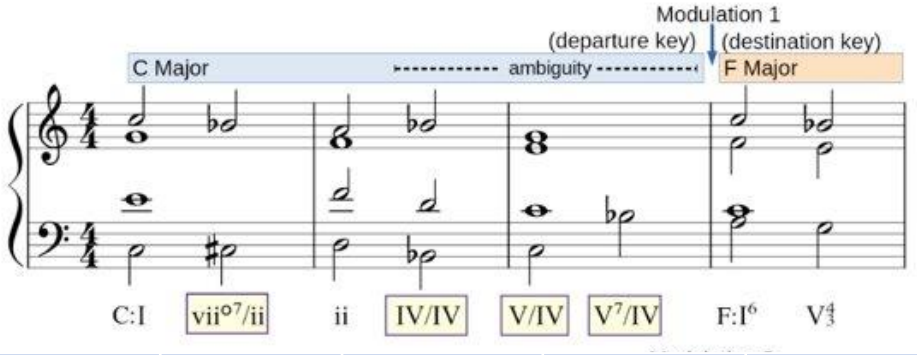
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

What do local keys coincide with?

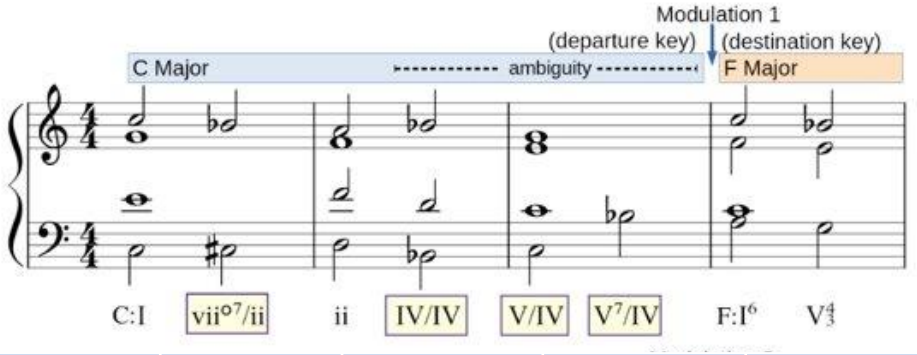
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Local keys vs. modulation

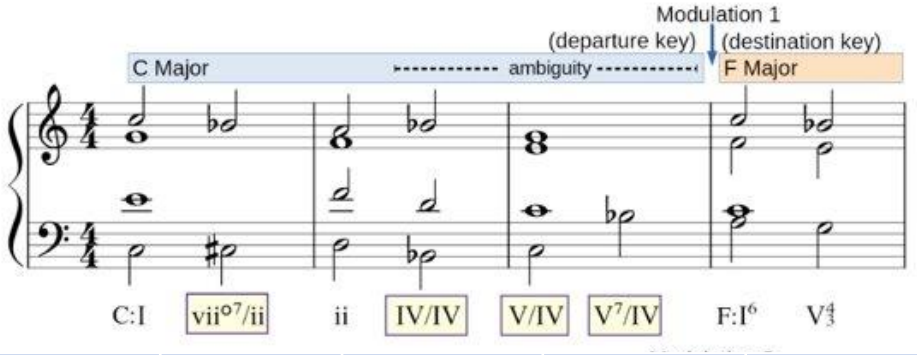
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Local keys vs. modulation

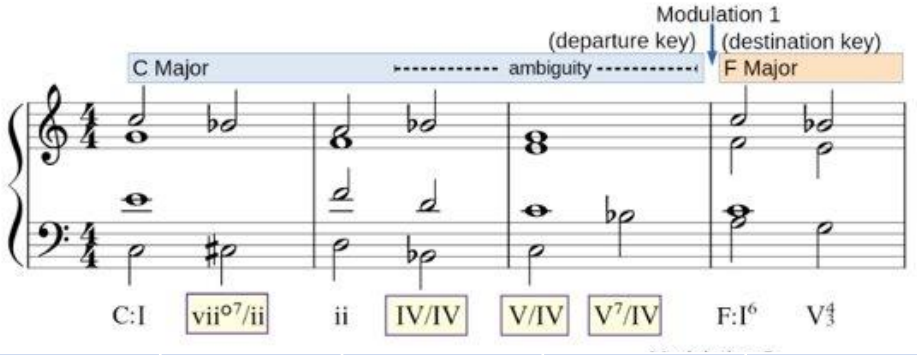
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Local keys vs. tonicization

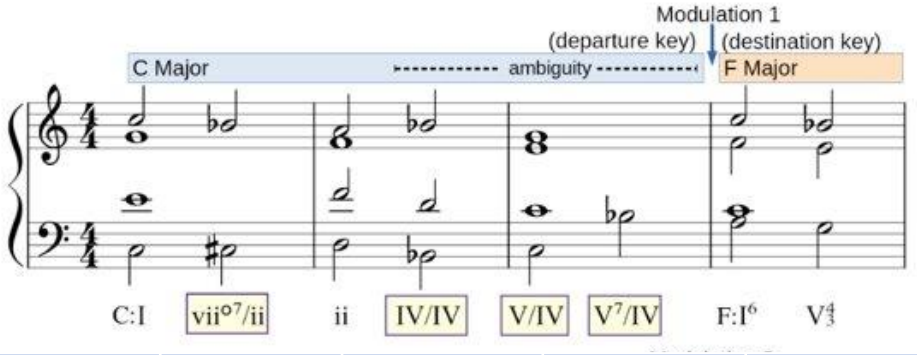
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Local keys vs. tonicization

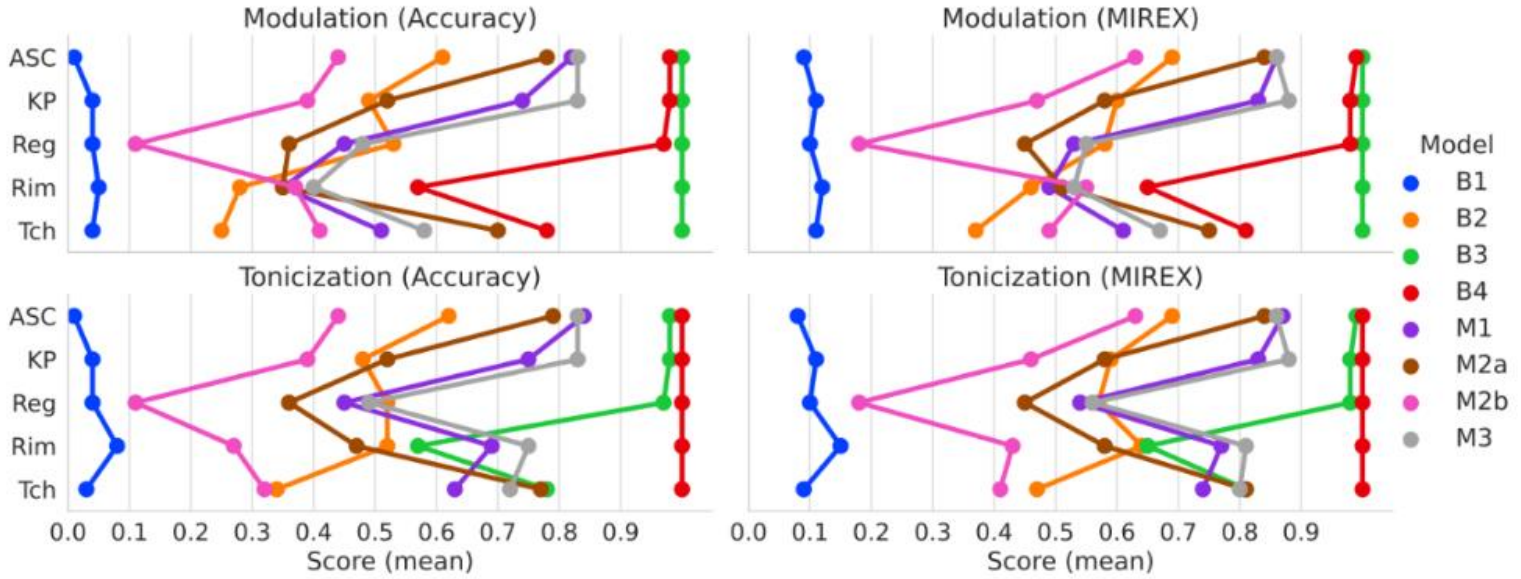
- Accuracy
- Mirex



Position	Annotation	Modulation (B3)	Tonicization (B4)	B1	B2	M1	M2a	M2b	M3
0	C:=>I	C	C	Bb	F	f	C	C	F
2	viio7/ii	C	d	a	F	F	C	C	F
4	ii	C	C	Bb	F	F	d	C	F
6	IV/IV	C	F	eb	F	F	d	C	F
8	V/IV	C	F	C#	F	F	d	C	F
10	V7/IV	C	F	f	F	F	d	C	F
12	F:=>I6	F	F	C#	F	F	d	C	F
14	V43	F	F	g	F	F	d	C	F
16	I	F	F	c	F	F	F	C	F
18	V2/V	F	C	C#	F	C	F	C	F
20	V6	F	F	C#	F	C	F	C	C
22	V	F	F	g	F	C	F	C	C
24	I6	F	F	b	F	C	F	F	F
26	V7/V	F	C	e	F	C	F	F	C
28	C:=>I	C	C	d	F	C	F	F	C

Results

- Random guess (B1) is the worst-performing model
- Global key guess (B2) is OK, but except in one textbook, it is behind the specialized local-key-estimation models
- Use of tonicization varies by composer
- When there are many tonicizations (Rim and Tch), models seem to coincide better with tonicizations than modulations (unexpected!)



Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.05	0.03	0.02	0.04	0.03
	Ton	0.05	0.03	0.02	0.07	0.05
B2	Mod	0.61	0.49	**0.53**	0.28	0.25
	Ton	0.62	0.48	**0.52**	0.52	0.34
M1	Mod	0.82	0.74	0.45	0.35	0.51
	Ton	**0.84**	0.75	0.45	0.69	0.63
M2a	Mod	0.78	0.52	0.36	0.35	**0.70**
	Ton	0.79	0.52	0.36	0.47	**0.77**
M2b	Mod	0.44	0.39	0.11	0.37	0.41
	Ton	0.44	0.39	0.11	0.27	0.32
M3	Mod	**0.83**	**0.83**	0.48	**0.40**	0.58
	Ton	0.83	**0.83**	0.49	**0.75**	0.72
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.98	0.98	0.97	0.57	0.78
B4	Mod	0.98	0.98	0.97	0.57	0.78
	Ton	1.00	1.00	1.00	1.00	1.00

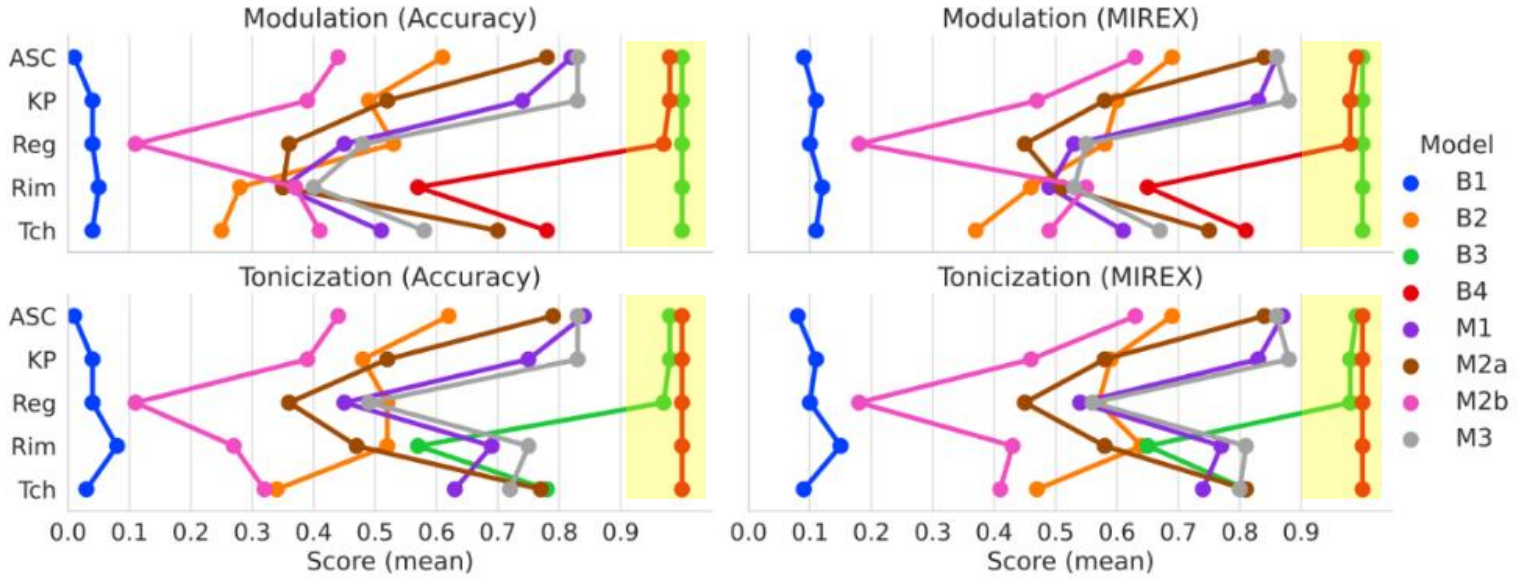
Accuracy weights

Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.09	0.10	0.08	0.11	0.10
	Ton	0.10	0.10	0.08	0.13	0.13
B2	Mod	0.69	0.60	**0.58**	0.46	0.37
	Ton	0.69	0.59	**0.57**	0.64	0.47
M1	Mod	**0.86**	0.83	0.53	0.49	0.61
	Ton	**0.87**	0.83	0.54	0.77	0.74
M2a	Mod	0.84	0.58	0.45	0.51	**0.75**
	Ton	0.84	0.58	0.45	0.58	**0.81**
M2b	Mod	0.63	0.47	0.18	**0.55**	0.49
	Ton	0.63	0.46	0.18	0.43	0.41
M3	Mod	**0.86**	**0.88**	0.55	0.53	0.67
	Ton	0.86	**0.88**	0.56	**0.81**	0.80
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.99	0.98	0.98	0.65	0.81
B4	Mod	0.99	0.98	0.98	0.65	0.81
	Ton	1.00	1.00	1.00	1.00	1.00

MIREX weights

Results

- Random guess (B1) is the worst-performing model
- Global key guess (B2) is OK, but except in one textbook, it is behind the specialized local-key-estimation models
- Use of tonicization varies by composer
- When there are many tonicizations (Rim and Tch), models seem to coincide better with tonicizations than modulations (unexpected!)



Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.05	0.03	0.02	0.04	0.03
	Ton	0.05	0.03	0.02	0.07	0.05
B2	Mod	0.61	0.49	**0.53**	0.28	0.25
	Ton	0.62	0.48	**0.52**	0.52	0.34
M1	Mod	0.82	0.74	0.45	0.35	0.51
	Ton	**0.84**	0.75	0.45	0.69	0.63
M2a	Mod	0.78	0.52	0.36	0.35	**0.70**
	Ton	0.79	0.52	0.36	0.47	**0.77**
M2b	Mod	0.44	0.39	0.11	0.37	0.41
	Ton	0.44	0.39	0.11	0.27	0.32
M3	Mod	**0.83**	**0.83**	0.48	**0.40**	0.58
	Ton	0.83	**0.83**	0.49	**0.75**	0.72
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.98	0.98	0.97	0.57	0.78
B4	Mod	0.98	0.98	0.97	0.57	0.78
	Ton	1.00	1.00	1.00	1.00	1.00

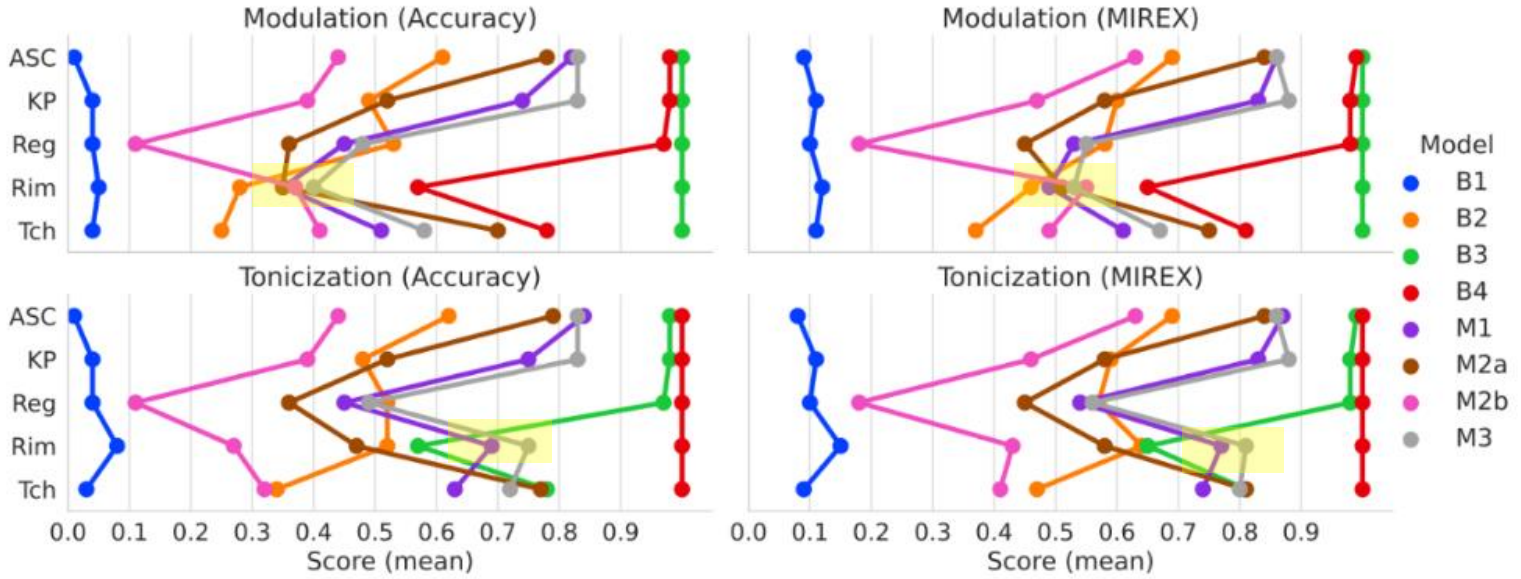
Accuracy weights

Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.09	0.10	0.08	0.11	0.10
	Ton	0.10	0.10	0.08	0.13	0.13
B2	Mod	0.69	0.60	**0.58**	0.46	0.37
	Ton	0.69	0.59	**0.57**	0.64	0.47
M1	Mod	**0.86**	0.83	0.53	0.49	0.61
	Ton	**0.87**	0.83	0.54	0.77	0.74
M2a	Mod	0.84	0.58	0.45	0.51	**0.75**
	Ton	0.84	0.58	0.45	0.58	**0.81**
M2b	Mod	0.63	0.47	0.18	**0.55**	0.49
	Ton	0.63	0.46	0.18	0.43	0.41
M3	Mod	**0.86**	**0.88**	0.55	0.53	0.67
	Ton	0.86	**0.88**	0.56	**0.81**	0.80
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.99	0.98	0.98	0.65	0.81
B4	Mod	0.99	0.98	0.98	0.65	0.81
	Ton	1.00	1.00	1.00	1.00	1.00

MIREX weights

Results

- Random guess (B1) is the worst-performing model
- Global key guess (B2) is OK, but except in one textbook, it is behind the specialized local-key-estimation models
- Use of tonicization varies by composer
- When there are many tonicizations (Rim and Tch), models seem to coincide better with tonicizations than modulations (unexpected!)



Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.05	0.03	0.02	0.04	0.03
	Ton	0.05	0.03	0.02	0.07	0.05
B2	Mod	0.61	0.49	**0.53**	0.28	0.25
	Ton	0.62	0.48	**0.52**	0.52	0.34
M1	Mod	0.82	0.74	0.45	0.35	0.51
	Ton	**0.84**	0.75	0.45	0.69	0.63
M2a	Mod	0.78	0.52	0.36	0.35	**0.70**
	Ton	0.79	0.52	0.36	0.47	**0.77**
M2b	Mod	0.44	0.39	0.11	0.37	0.41
	Ton	0.44	0.39	0.11	0.27	0.32
M3	Mod	**0.83**	**0.83**	0.48	**0.40**	0.58
	Ton	0.83	**0.83**	0.49	**0.75**	0.72
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.98	0.98	0.97	0.57	0.78
B4	Mod	0.98	0.98	0.97	0.57	0.78
	Ton	1.00	1.00	1.00	1.00	1.00

Accuracy weights

Model	Task	ASC	KP	Reg	Rim	Tch
B1	Mod	0.09	0.10	0.08	0.11	0.10
	Ton	0.10	0.10	0.08	0.13	0.13
B2	Mod	0.69	0.60	**0.58**	0.46	0.37
	Ton	0.69	0.59	**0.57**	0.64	0.47
M1	Mod	**0.86**	0.83	0.53	0.49	0.61
	Ton	**0.87**	0.83	0.54	0.77	0.74
M2a	Mod	0.84	0.58	0.45	0.51	**0.75**
	Ton	0.84	0.58	0.45	0.58	**0.81**
M2b	Mod	0.63	0.47	0.18	**0.55**	0.49
	Ton	0.63	0.46	0.18	0.43	0.41
M3	Mod	**0.86**	**0.88**	0.55	0.53	0.67
	Ton	0.86	**0.88**	0.56	**0.81**	0.80
B3	Mod	1.00	1.00	1.00	1.00	1.00
	Ton	0.99	0.98	0.98	0.65	0.81
B4	Mod	0.99	0.98	0.98	0.65	0.81
	Ton	1.00	1.00	1.00	1.00	1.00

MIREX weights

Conclusions

- Investigate relationship between local keys, modulations, and tonicizations
- Encoding modulation and tonicization ground truth in a dataset with 201 excerpts of music
- Method for comparing local keys to modulations and tonicizations
- Use of tonicizations among theorists varies considerably
- Generally, local-key-estimation models seem to coincide more with tonicizations

Thank you !

Dataset is available here :
https://github.com/DDMAL/key_modulation_dataset