

# What's a word in harmony?

Can Statistical Language Models be used for the  
Analysis of Harmonic Progressions?

Matthias Mauch  
Daniel Müllensiefen  
Geraint Wiggins  
Simon Dixon

Good afternoon.  
– title is “Can Statistical...” but really we’re asking “What’s a ...”  
– let me clarify this  
-->

# What's this about?

- ▶ Find a symbolic **harmony feature**  
whose type sparseness resembles that of types in natural language.
- ▶ Results Sneak Preview
  - ▶ Single chords don't do the job.
  - ▶ Longer chord sequences resemble words in language more closely.

Mauch, Müllensiefen, Wiggins, Dixon

# Motivation: Music IR

- ▶ example 1 :  
“find similar song given chord transcriptions” (Query By Example)
- ▶ latent semantic analysis
- ▶ example 2:  
“extract chords from complex audio”
- ▶ Chords sequential (like words)
- ▶ HMM with chords as states

Mauch, Müllensiefen, Wiggins, Dixon

# Music IR: audio to labels

- ▶ example 3:
  - “extract chords from complex audio”
  - ▶ Chords sequential (like words)
  - ▶ NLP methods work so well ...
  - ▶ HMM with chords as states

Mauch, Müllensiefen, Wiggins, Dixon

# Music IR: audio to labels

- ▶ example 3:

“extract chords from complex audio”

- ▶ Chords sequential (like words)  
Alternative to
- ▶ NLP methods work so well ...  
chord
- ▶ HMM with chords as states  
as basic unit?

Mauch, Müllensiefen, Wiggins, Dixon

# Symbolic chord and text data

## chords

### Community Corpus

2548 songs,  
195,874 chords

### Automatic Corpus

2592 songs,  
294,264 chords

## text

### Brown Corpus

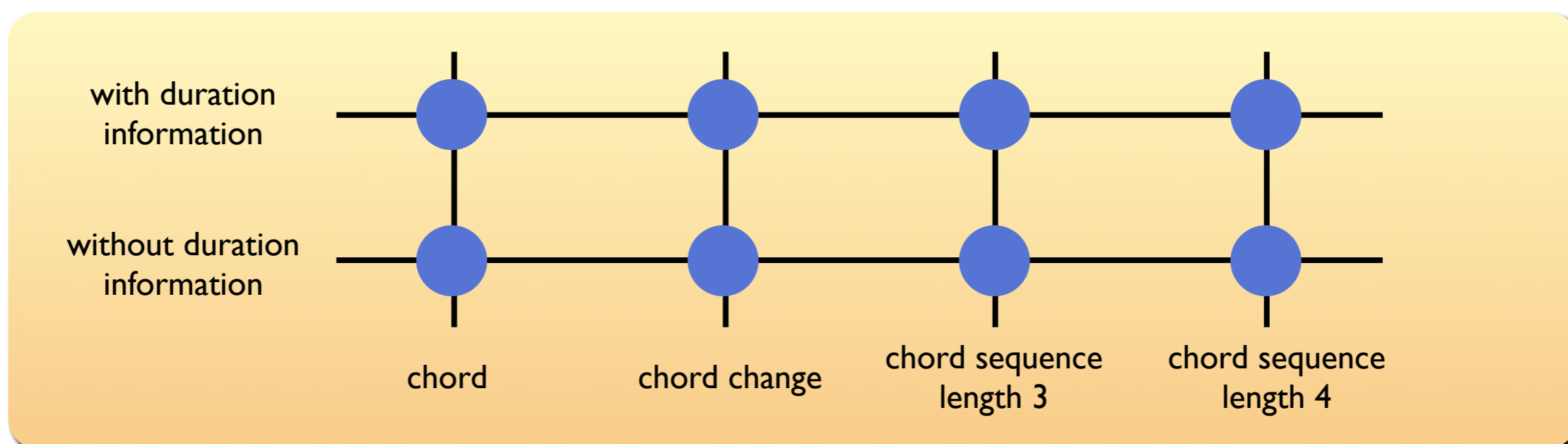
1,006,770 tokens (words)

Mauch, Müllensiefen, Wiggins, Dixon

- Community Corpus: (1) chord “lead sheets” shared by users of the commercial software “Band in a Box”, (2) checked by relistening via the program.
- Automatic Corpus: (1) chords automatically extracted from MIDI.

# Choice of harmonic feature

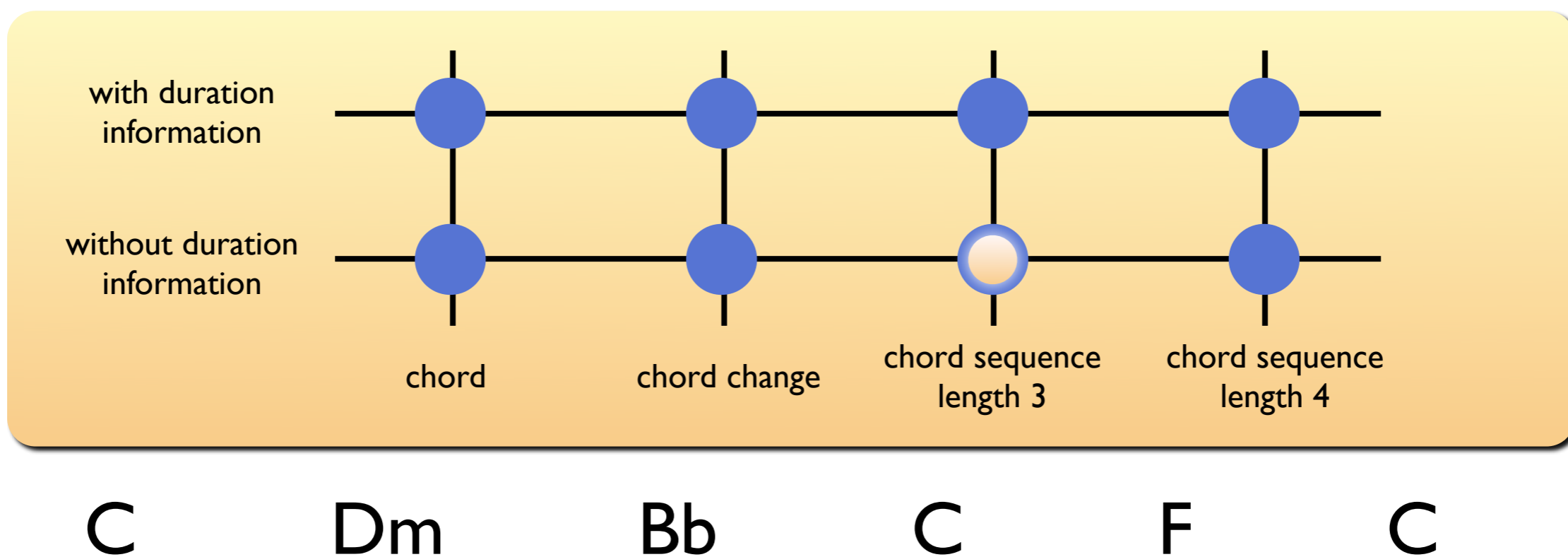
- ▶ compare word distribution to **chord sequence** distribution
- ▶ 8 kinds of chord sequences



Mauch, Müllensiefen, Wiggins, Dixon

- from the chord data we construct 8 different kinds of harmonic feature:
- chord sequences
- different lengths (ranging from a single chord to sequences of ...)
- another parameter is: consider harmonic rhythm or not (by considering duration info or not)
- Let me quickly show you some examples for such chord sequence types ....
- chord sequences overlap

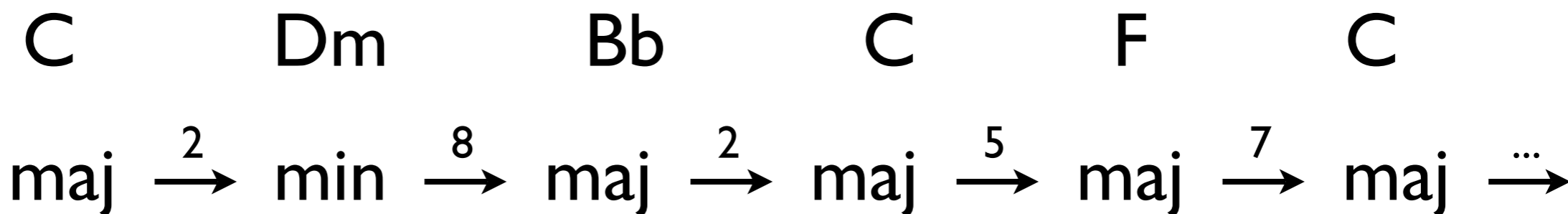
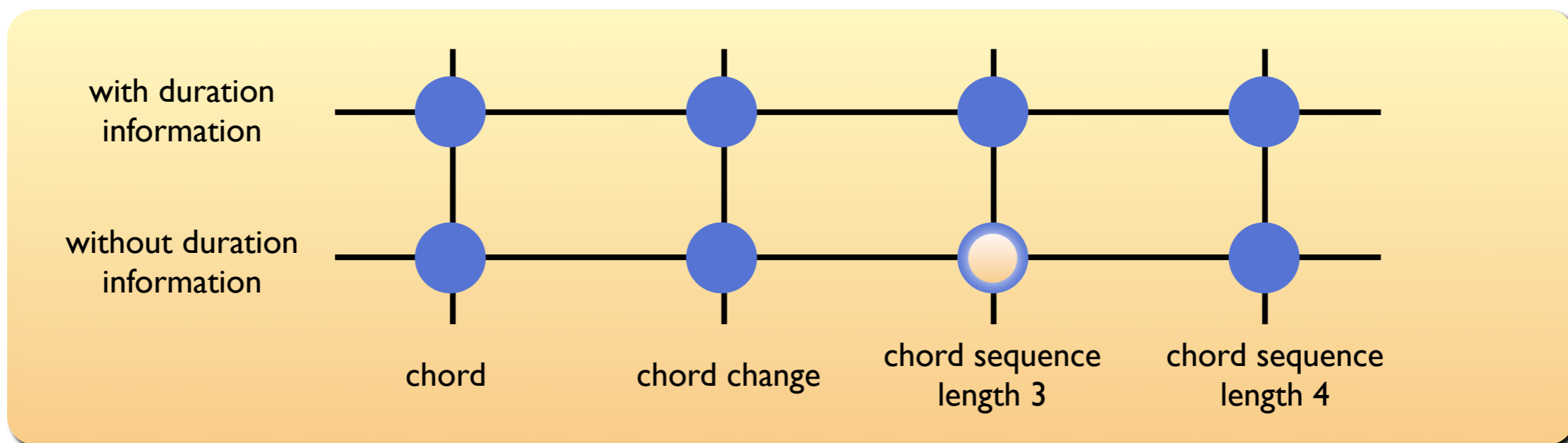
# Sequence extraction example



Mauch, Müllensiefen, Wiggins, Dixon

- explain for example setting “without duration information”, “chord sequence length 3”
- representation as chords with relative root distances in semitones
- sequences obtained by a sliding window
- benefit: key-independent, would be the same in a different key

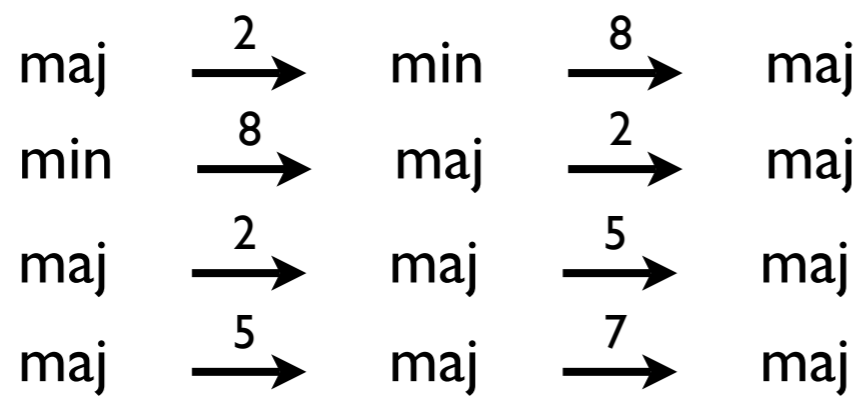
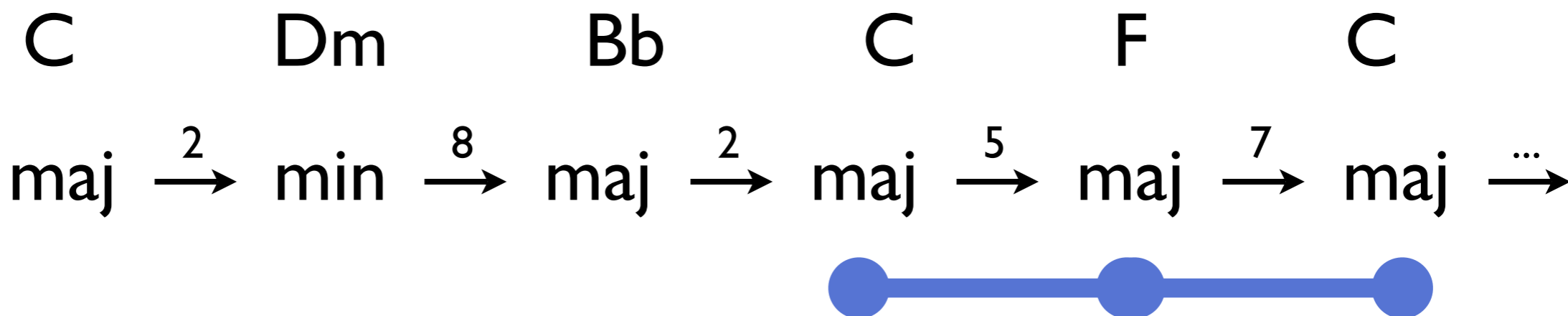
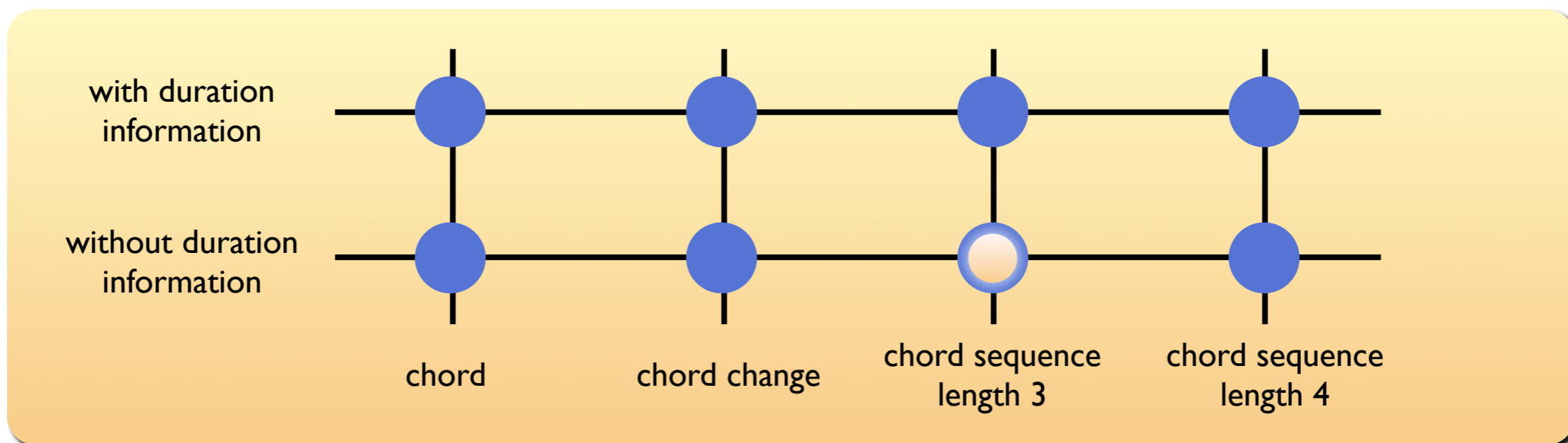
# Sequence extraction example



Mauch, Müllensiefen, Wiggins, Dixon

- explain for example setting “without duration information”, “chord sequence length 3”
- representation as chords with relative root distances in semitones
- sequences obtained by a sliding window
- benefit: key-independent, would be the same in a different key

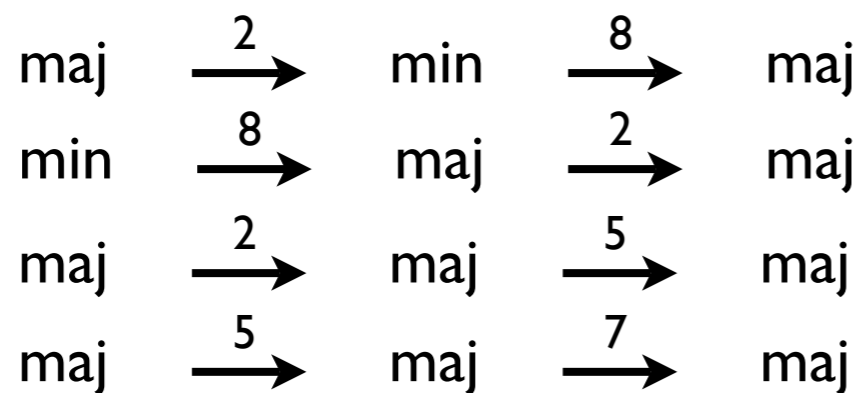
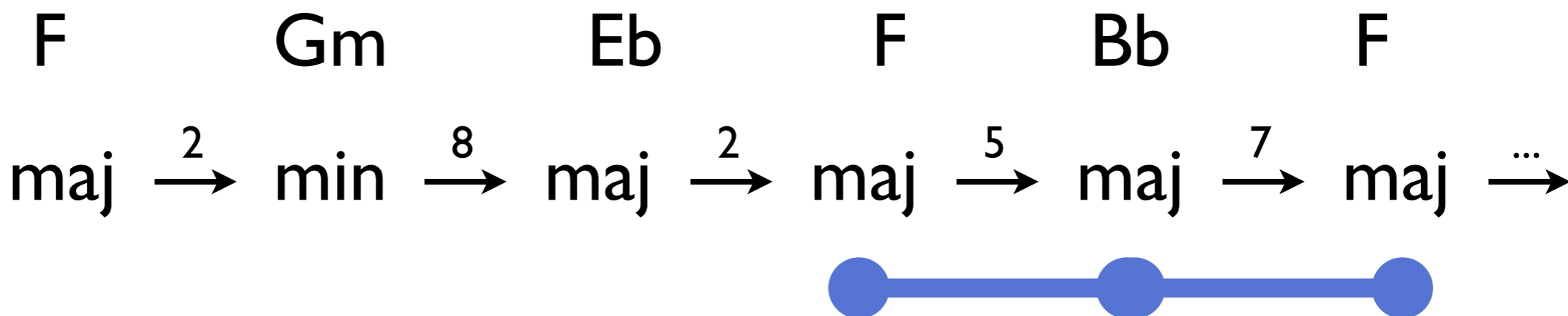
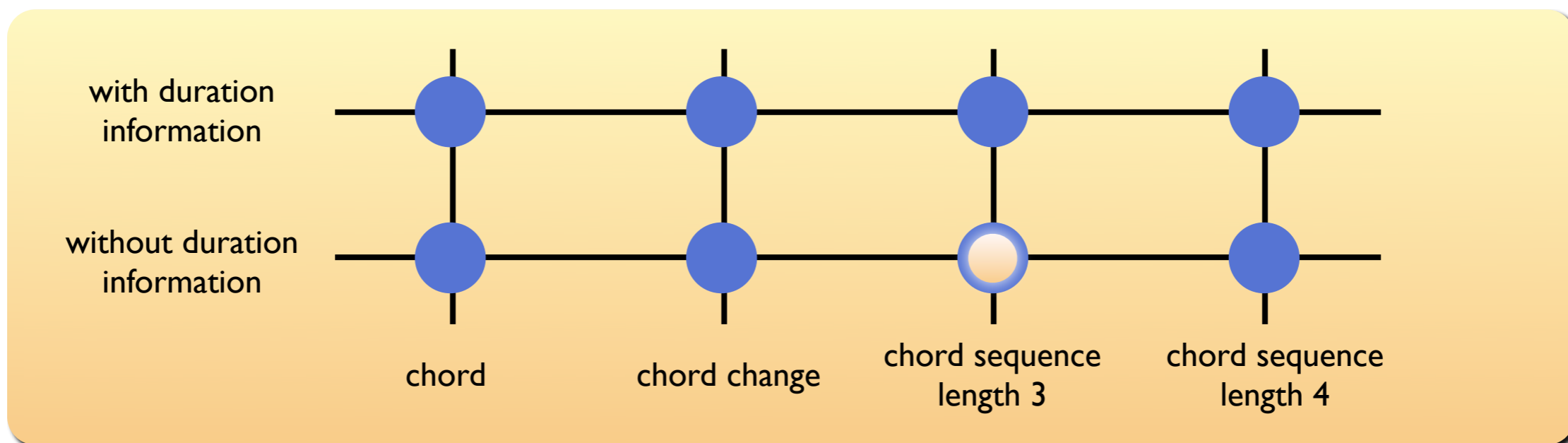
# Sequence extraction example



Mauch, Müllensiefen, Wiggins, Dixon

- explain for example setting “without duration information”, “chord sequence length 3”
- representation as chords with relative root distances in semitones
- sequences obtained by a sliding window
- benefit: key-independent, would be the same in a different key

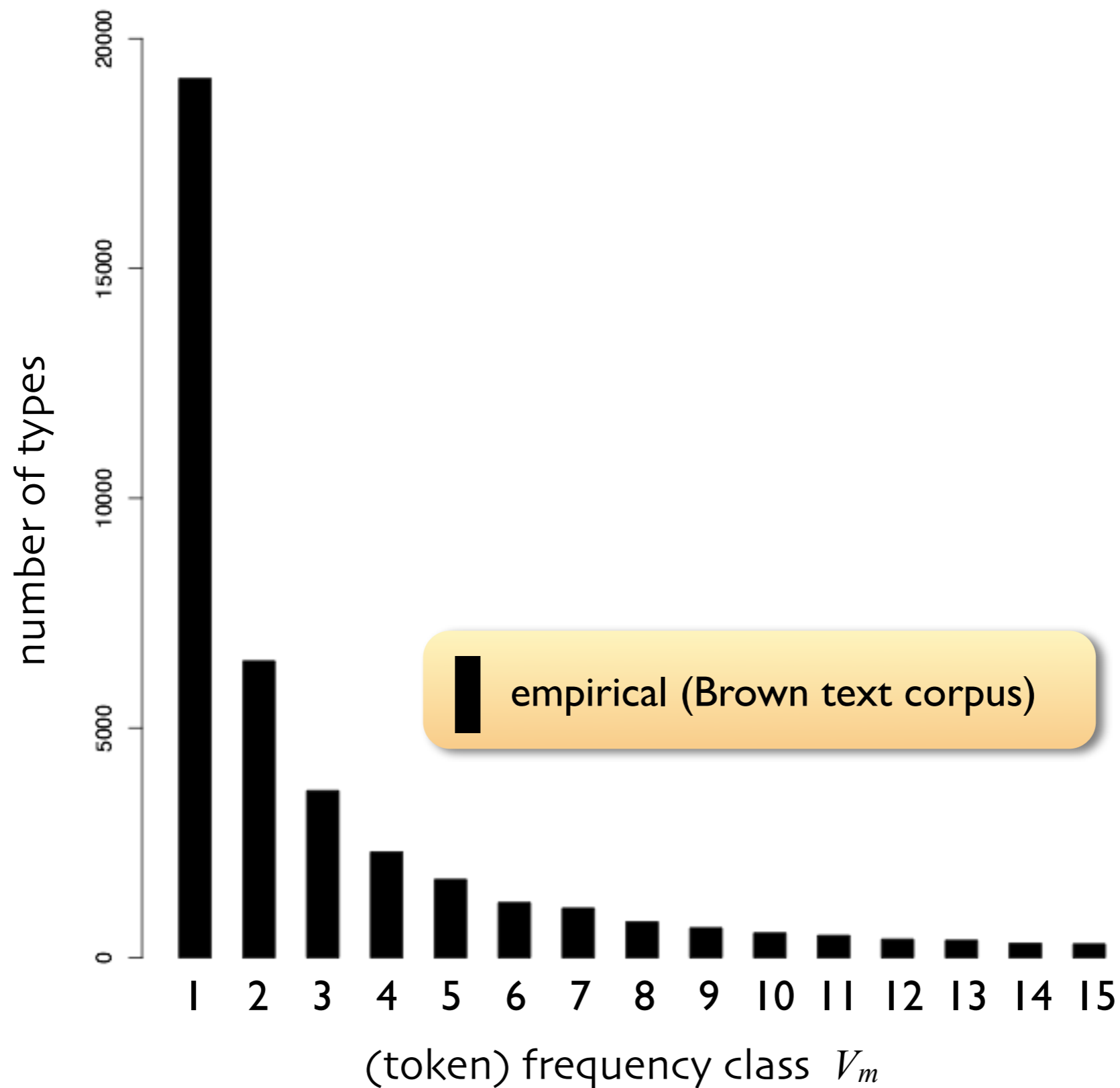
# Sequence extraction example



Mauch, Müllensiefen, Wiggins, Dixon

- explain for example setting “without duration information”, “chord sequence length 3”
- representation as chords with relative root distances in semitones
- sequences obtained by a sliding window
- benefit: key-independent, would be the same in a different key

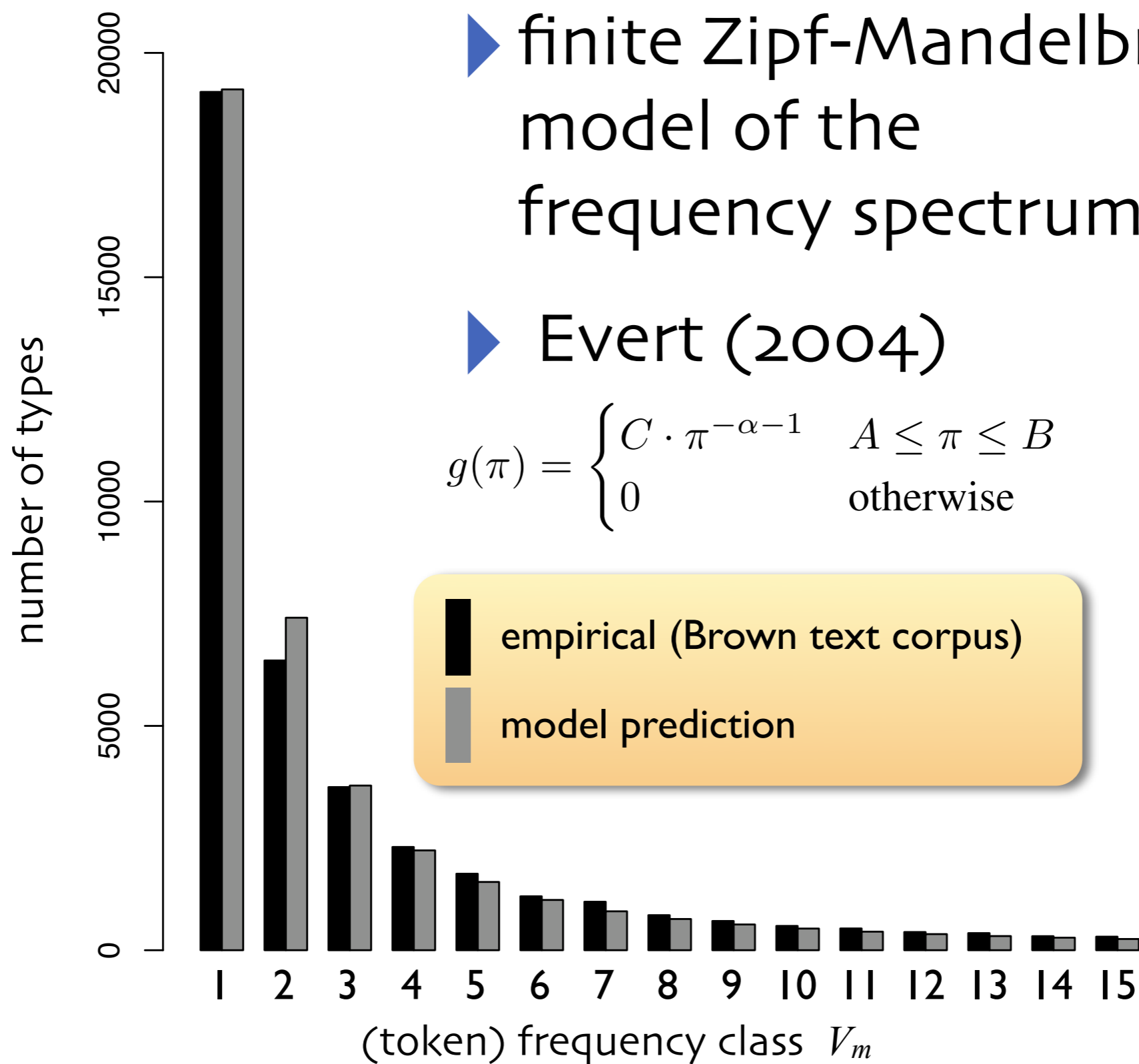
# Frequency Spectrum



Mauch, Müllensiefen, Wiggins, Dixon

- what we base our comparison on is the Frequency Spectrum:
- this one: text corpus example
  - “most words appear only once (frequency class  $V_{\{1\}}$ )”
  - smooth shape, can be modeled by a parametric model (next slide)

# Zipf-Mandelbrot Model



▶ finite Zipf-Mandelbrot model of the frequency spectrum

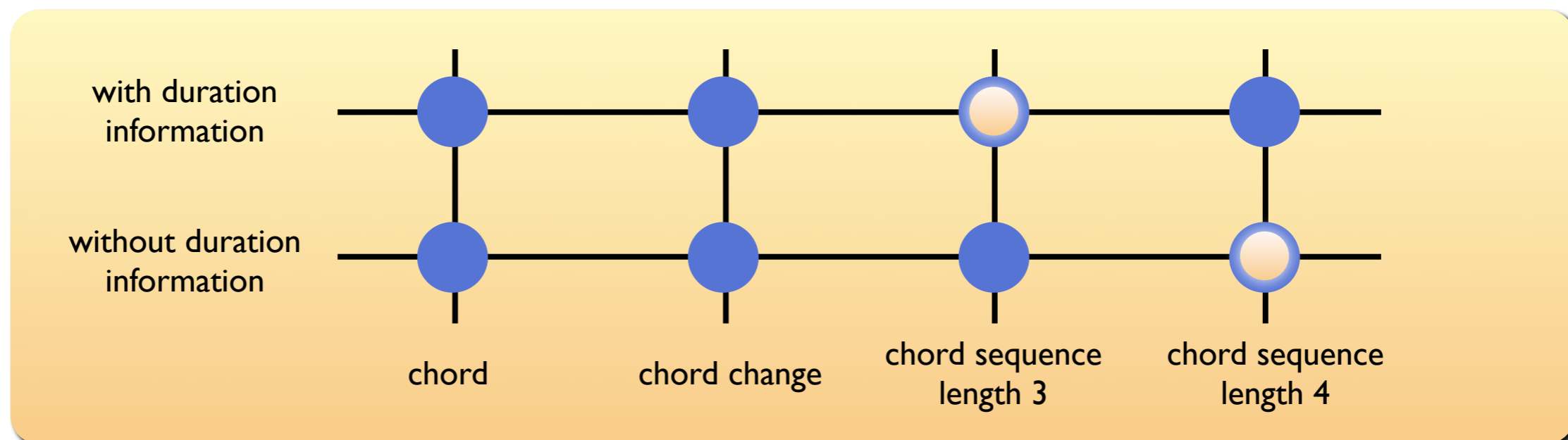
▶ Evert (2004)

$$g(\pi) = \begin{cases} C \cdot \pi^{-\alpha-1} & A \leq \pi \leq B \\ 0 & \text{otherwise} \end{cases}$$

Mauch, Müllensiefen, Wiggins, Dixon

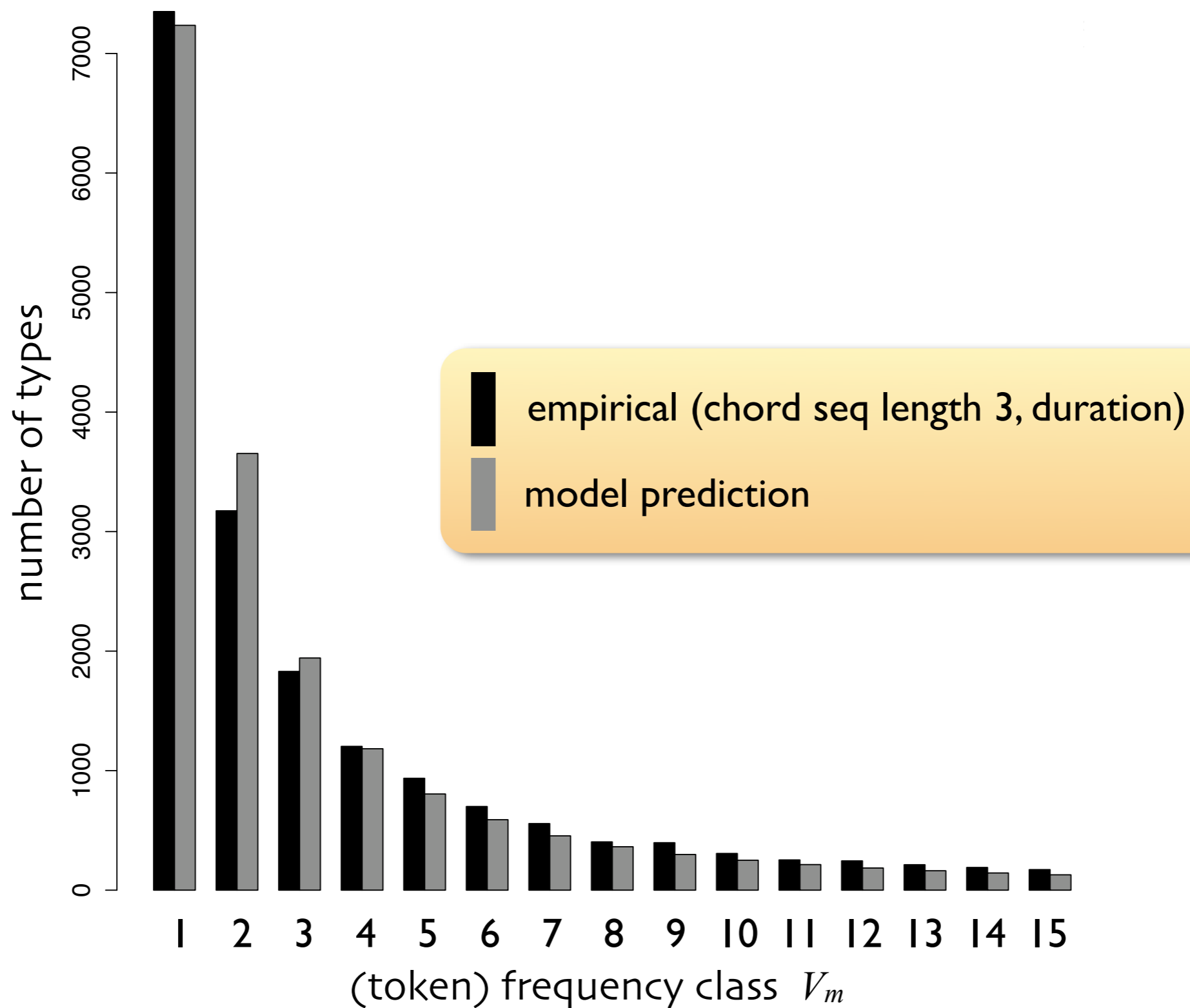
# Best-fitting Harmonic Elements

- ▶ settings closest to text corpus
  - ▶ length 3 (metric duration)
  - ▶ length 4 (no duration)



Mauch, Müllensiefen, Wiggins, Dixon

# Chord Sequence Distribution



Mauch, Müllensiefen, Wiggins, Dixon

# Interpretation of Results

- ▶ Single chords don't do the job.
- ▶ Longer chord sequences resemble words in language more closely.
- ▶ Even more so if considering durations.

Mauch, Müllensiefen, Wiggins, Dixon

# Future Work

- ▶ different harmonic elements
  - ▶ root progression
  - ▶ degree (with respect to key)
  - ▶ variable length chord progressions  
(determine using collocation measure)
- ▶ different language elements
  - ▶ letter sequences
  - ▶ other languages? Chinese?

Mauch, Müllensiefen, Wiggins, Dixon

# Thanks

▶ ... looking forward to your suggestions!

Mauch, Müllensiefen, Wiggins, Dixon