

# **Extended playing techniques: The next milestone in musical instrument recognition**

**Vincent Lostanlen, New York University**

**Joakim Andén, Flatiron Institute**

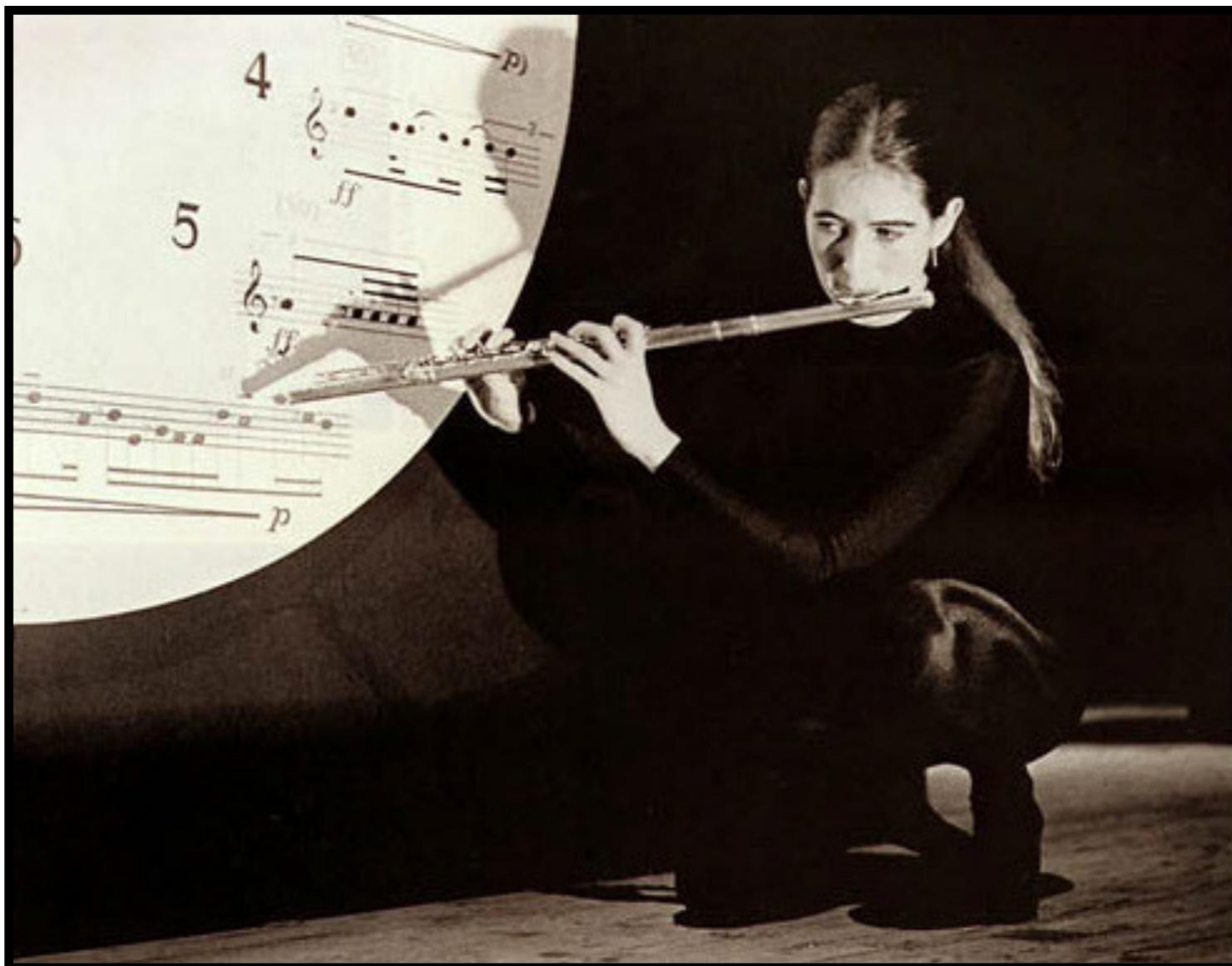
**Mathieu Lagrange, École Centrale de Nantes, CNRS**



Paris, Sept 2018. This work is supported by the ERC, the NSF, and Google.

# In a nutshell

Our definition of **timbre** requires a more systematic inclusion of **performance effects**.



*Kathinka  
Pasveer*

# Context: spectralist music creation

Analysis   Explorer   Filters   Orchestra   Solutions   Maquette   Quit

Orchestra

Preset: Modern a 4

- 2 Fl+Picc
- 2 Fl
- 3 Ob
- 1 Ob+EH
- 1 CIBb+CIEb+CIA
- 2 CIBb+CIA
- 1 CIBb+BCIBb
- 3 Bn
- 1 Bn+Cbn
- 4 Hn
- 3 TpC
- 3 TTbn
- 1 BTb
- 1 Hp
- 2 Vns
- 1 Vas
- 1 Vcs
- 

Stage mixer:

Reset Positioning and Gains   Reset Gains

Apply Orchestra

[Eslin 2014]

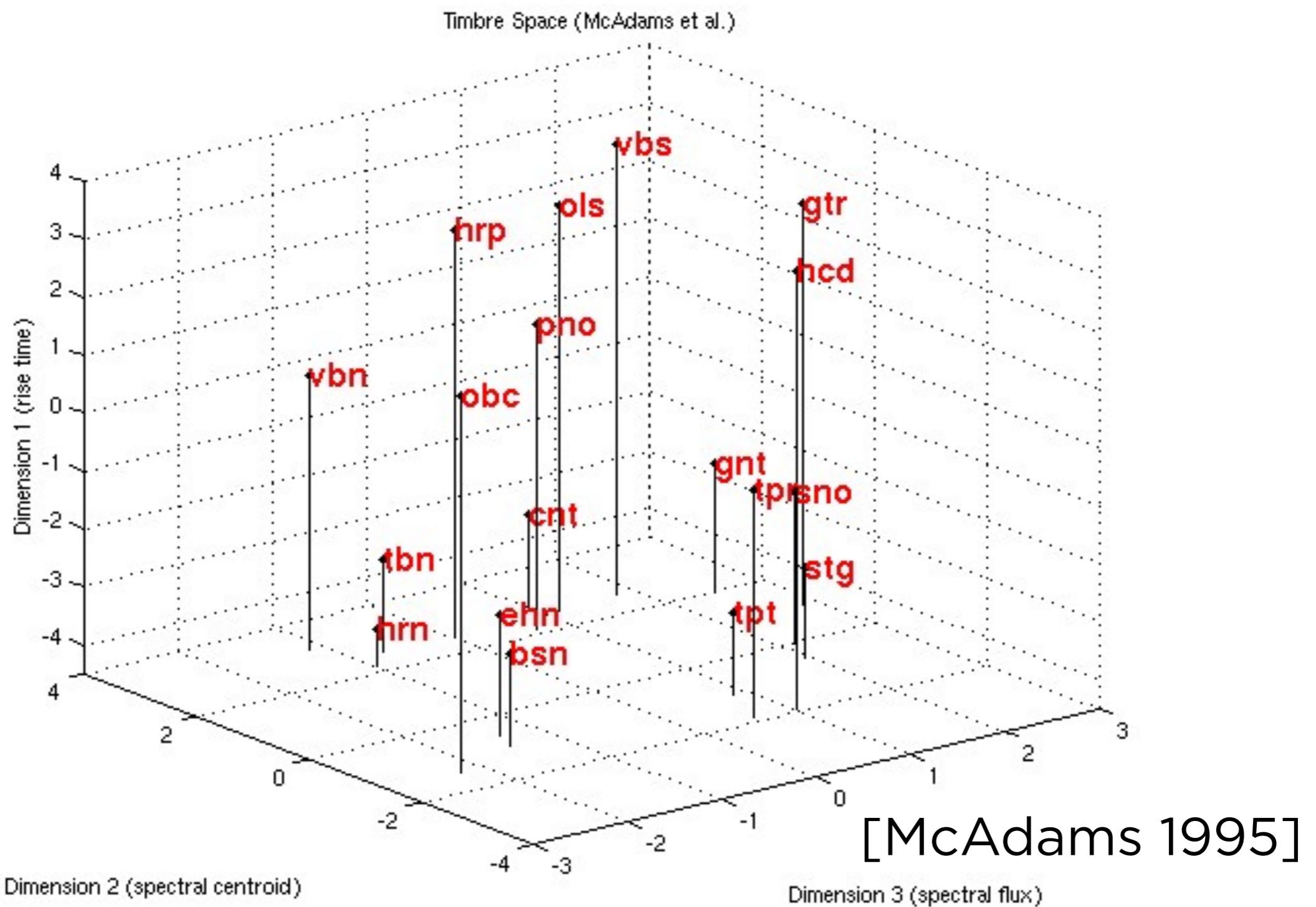
# *What are musical instruments?*

## the organology definition:

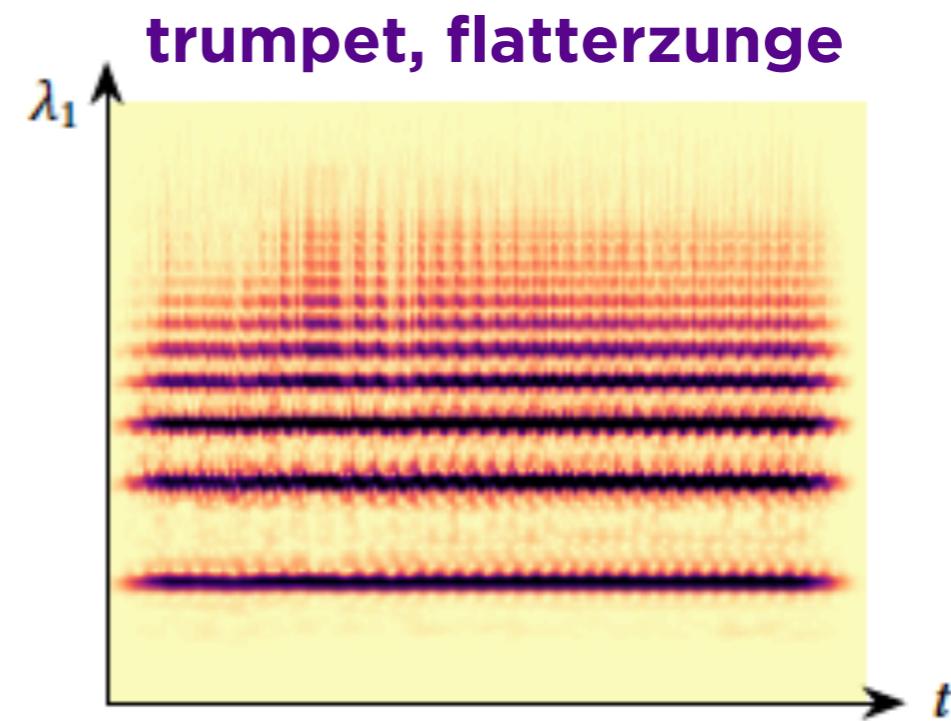
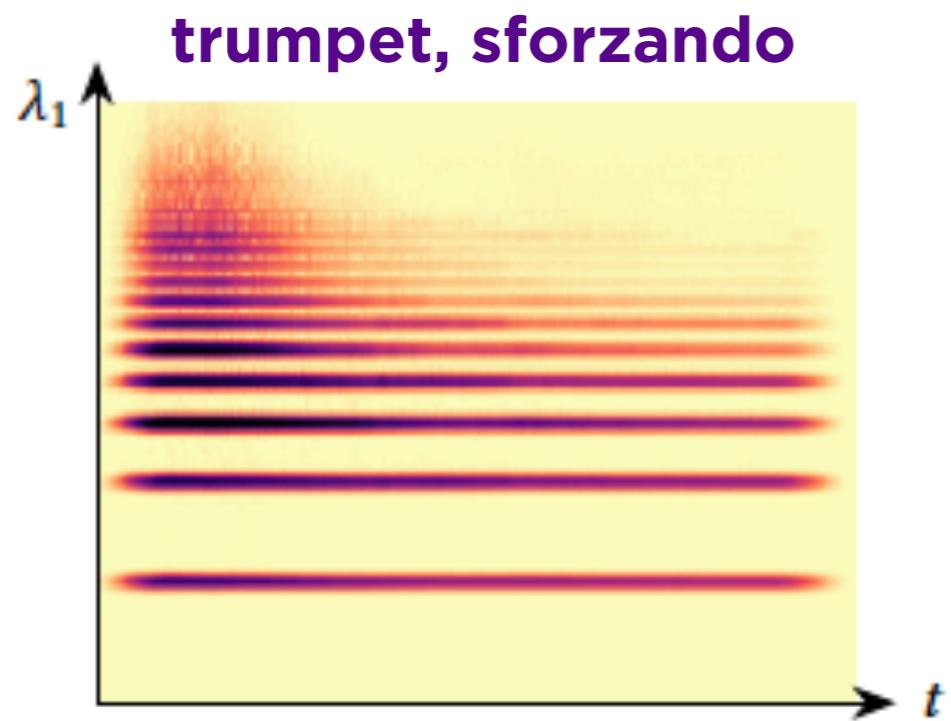
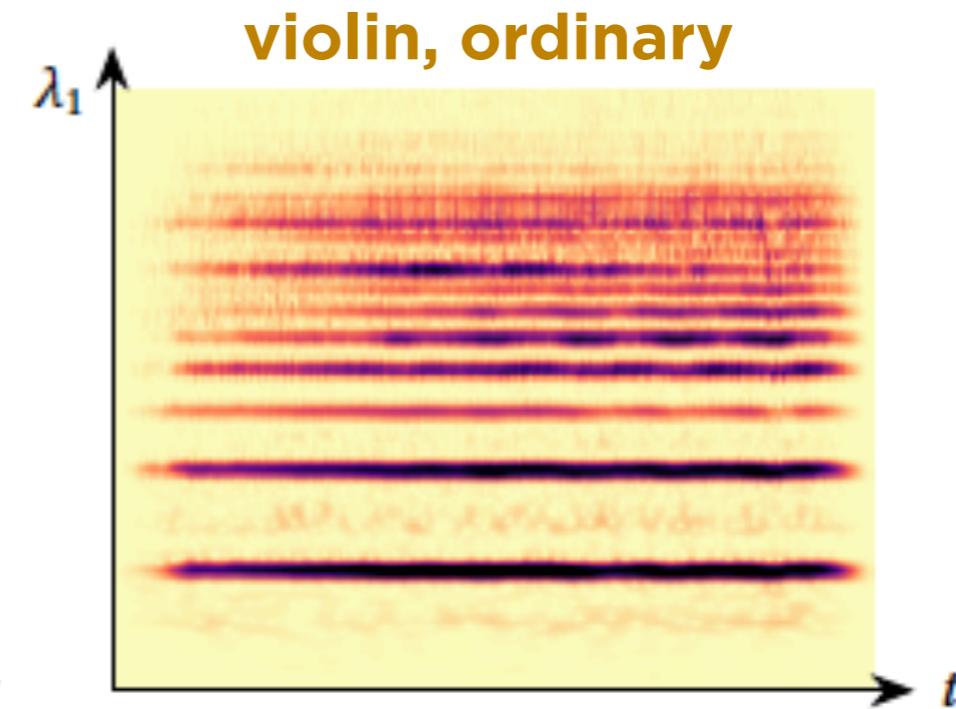
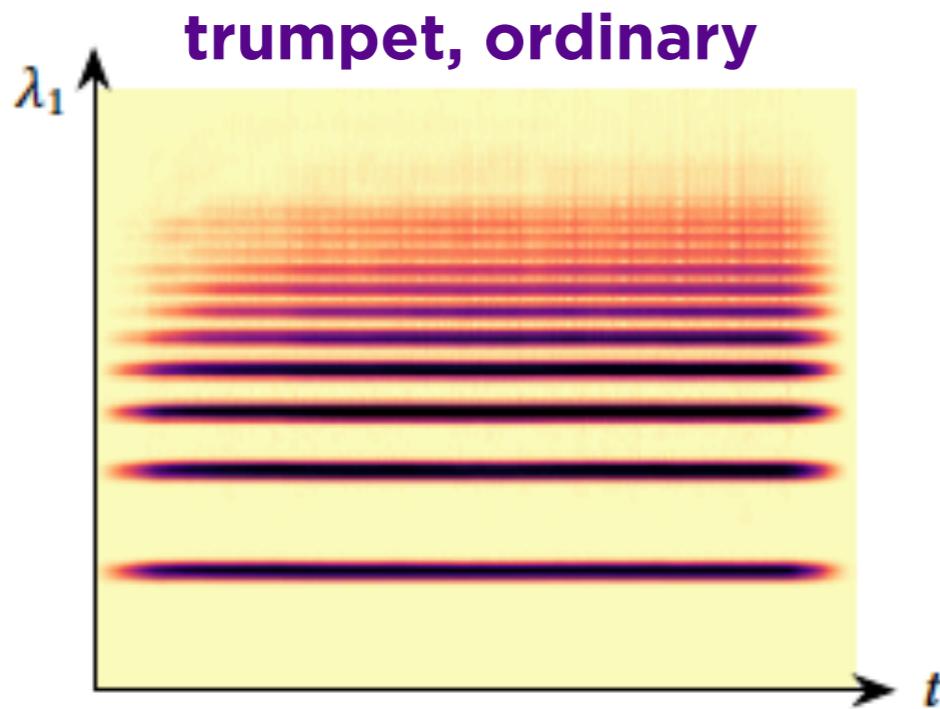


# *What are musical instruments?*

## the music psychology definition:



Problem: none of the two definitions account for variability in instrumental playing techniques (IPT).



Yet, non-ordinary playing techniques are frequent

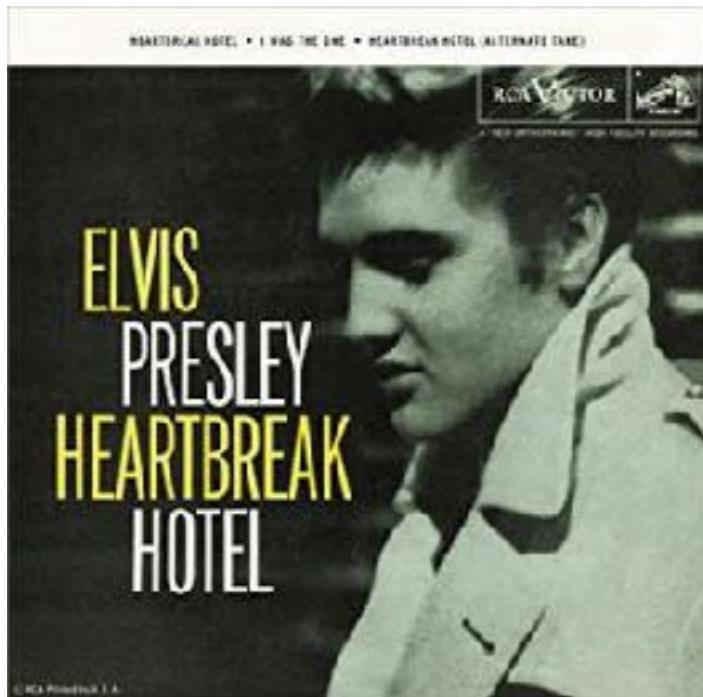
**clarinet slide**



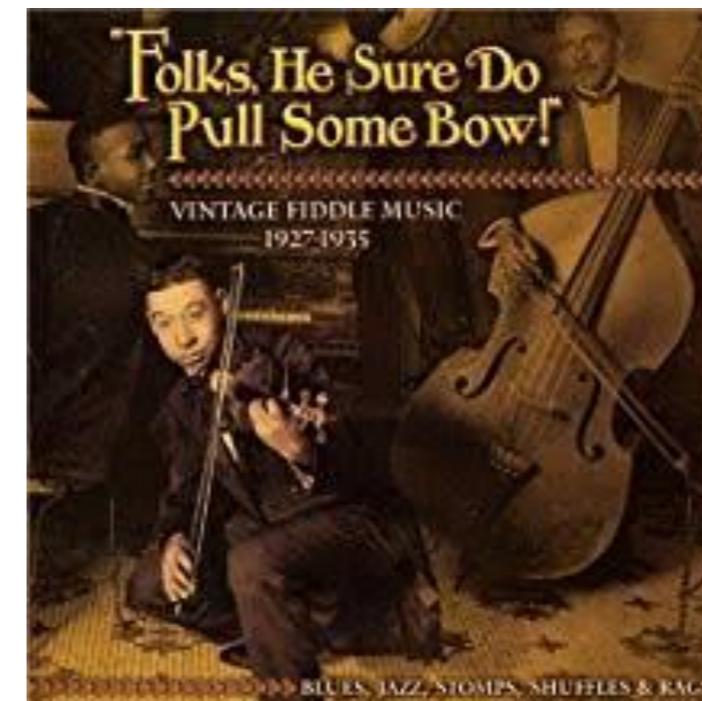
**saxophone growl**

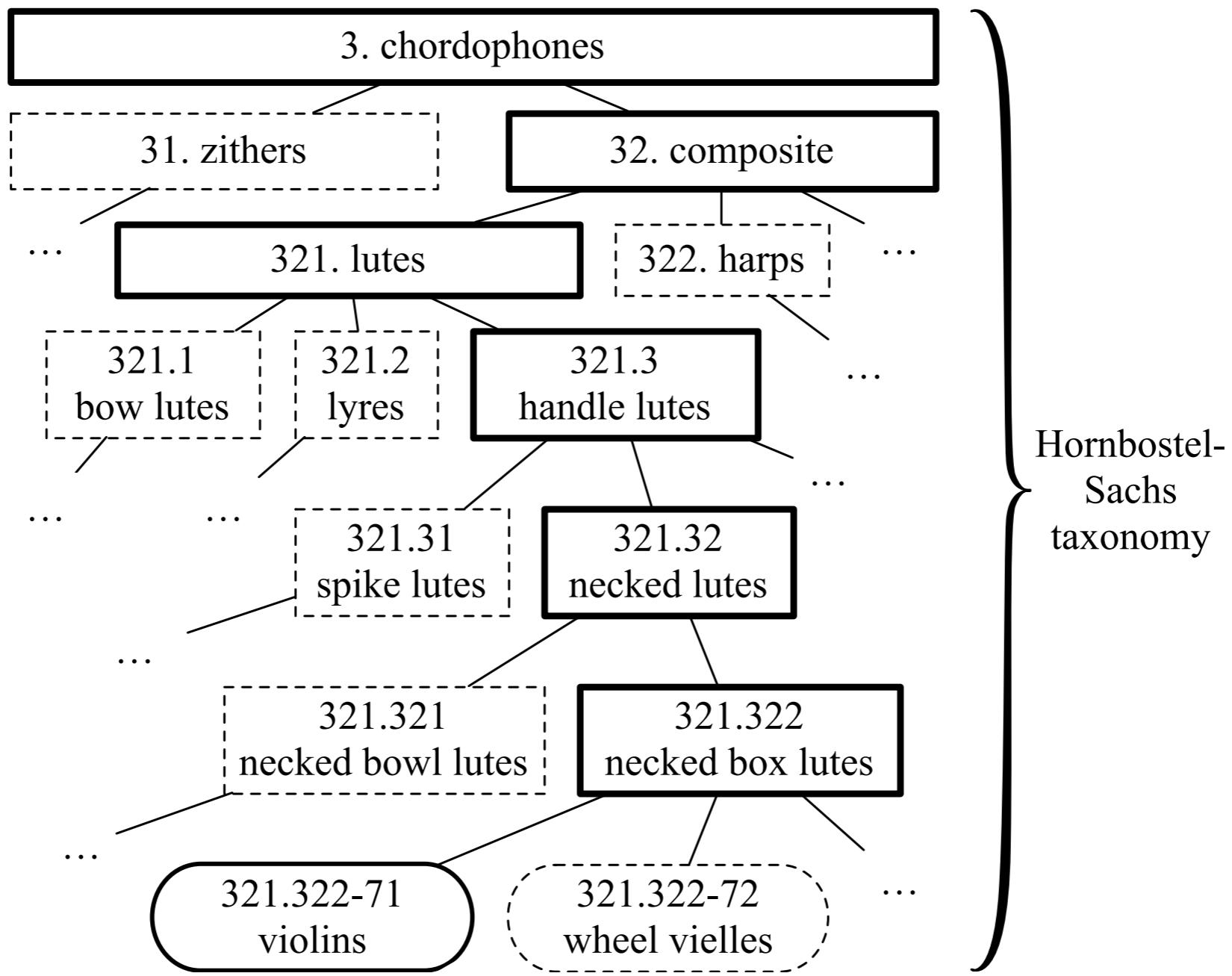


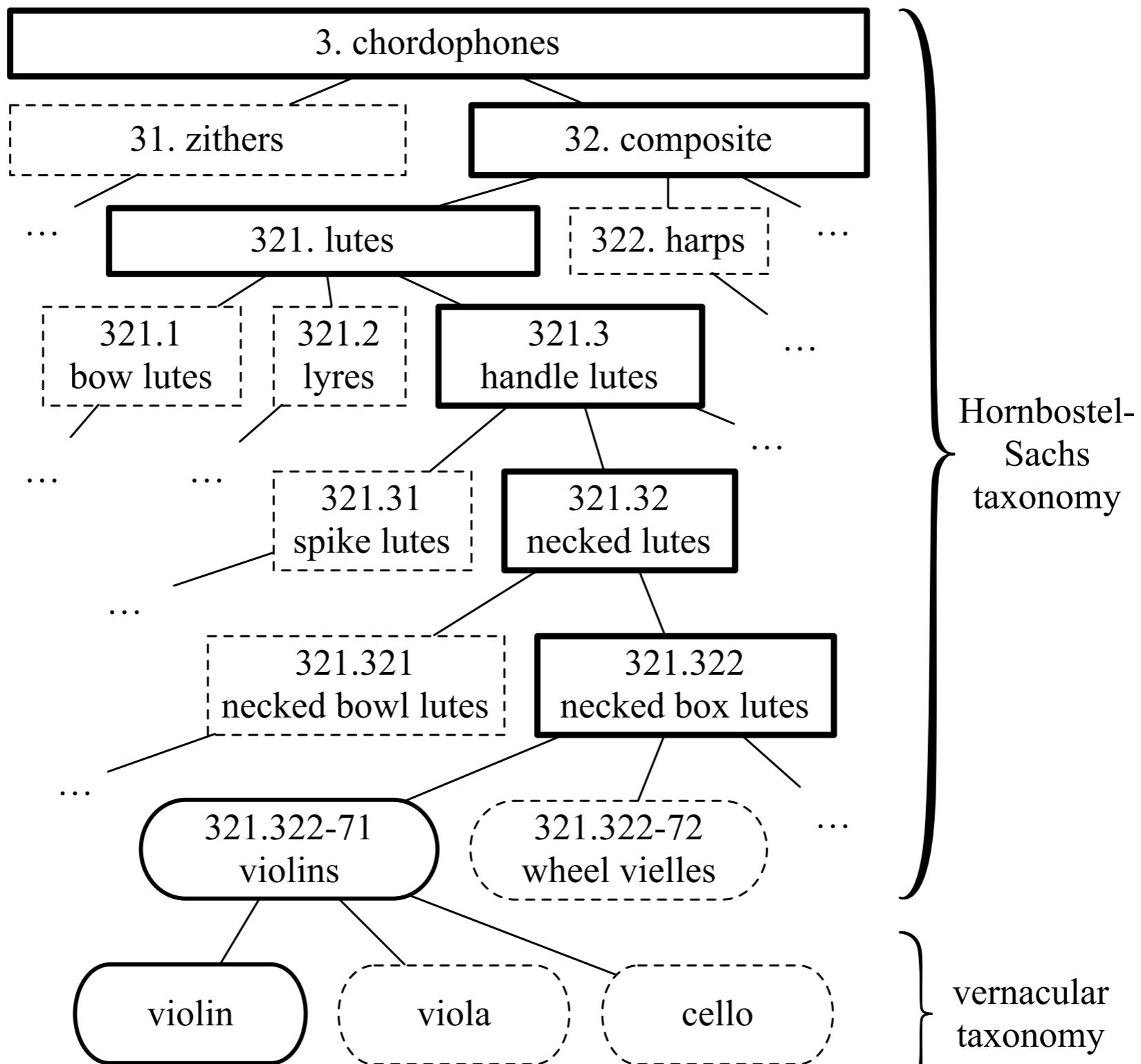
**bass slap**



**violin shuffle stroke**

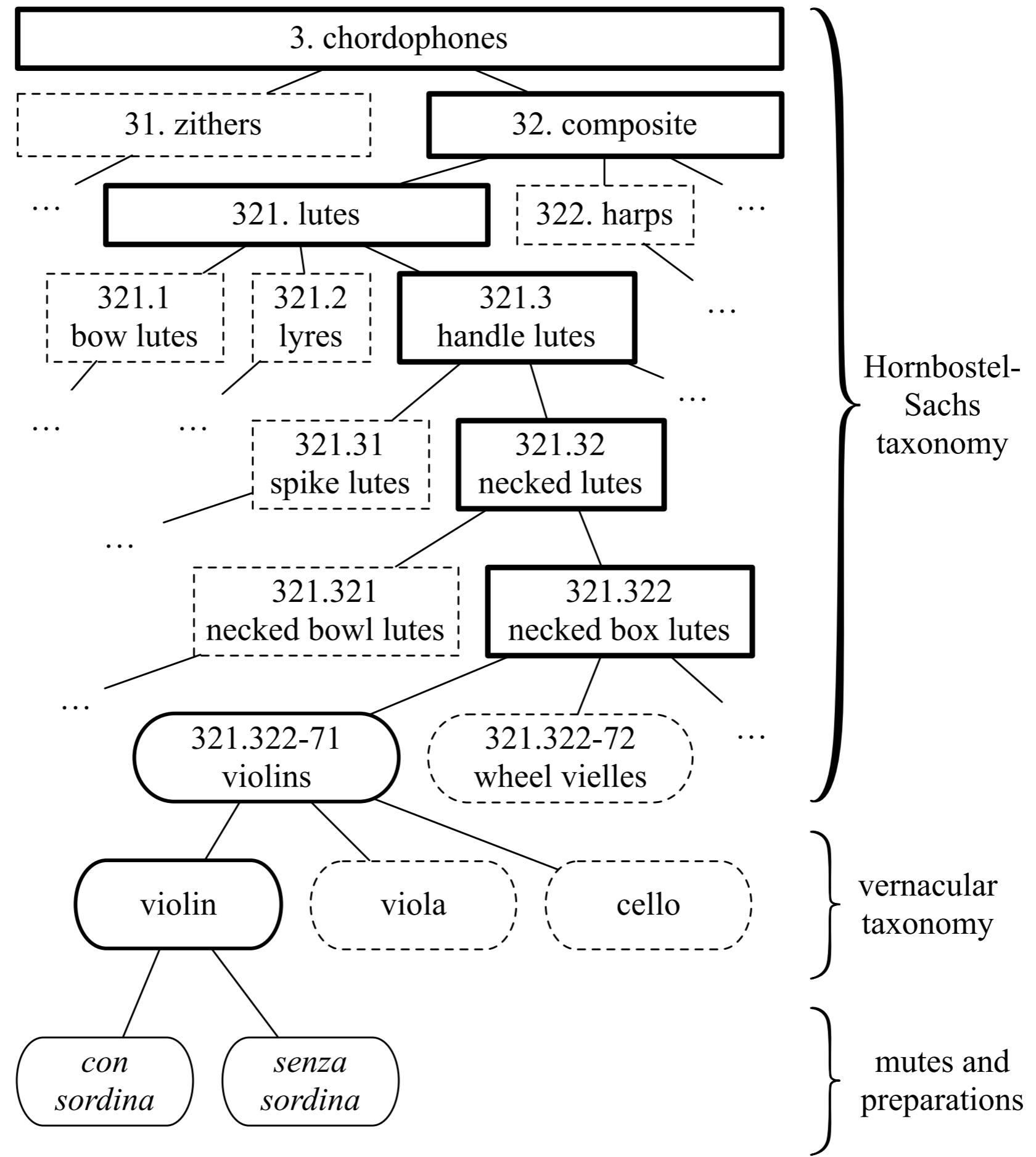






Hornbostel-Sachs taxonomy

vernacular taxonomy



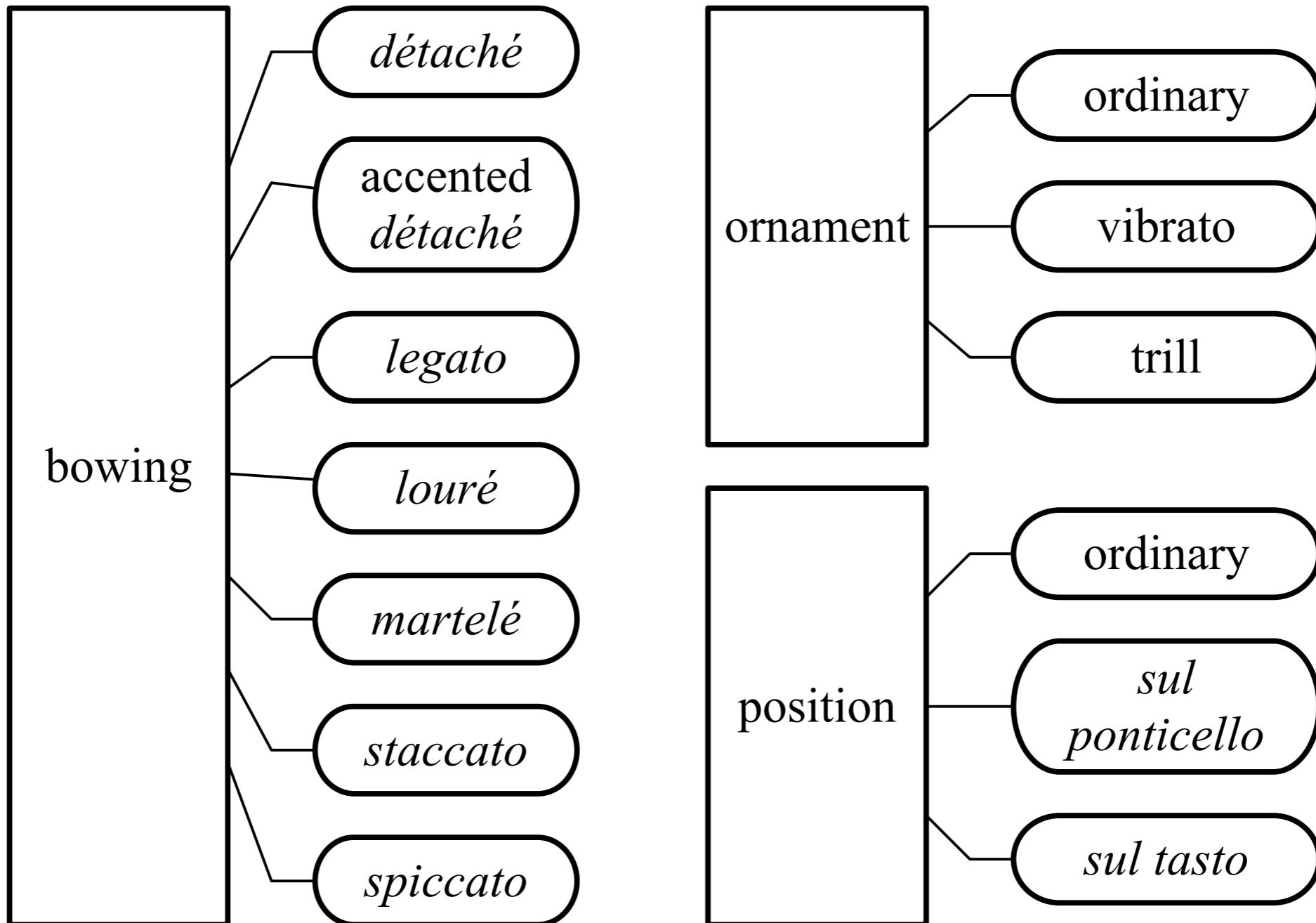
- Musical instruments, as inert objects, may abide to a **hierarchical taxonomy**.
- Yet, instrumental playing techniques form a loose **folksonomy of namespaces**.

An MIR appeal to musicological expertise:  
*How to reconcile organology and “chironomics”?*

We (tentatively) formulate the problem as **query-by-example** similarity retrieval, with a precision@5 evaluation metric.

# Example: namespaces of the violin

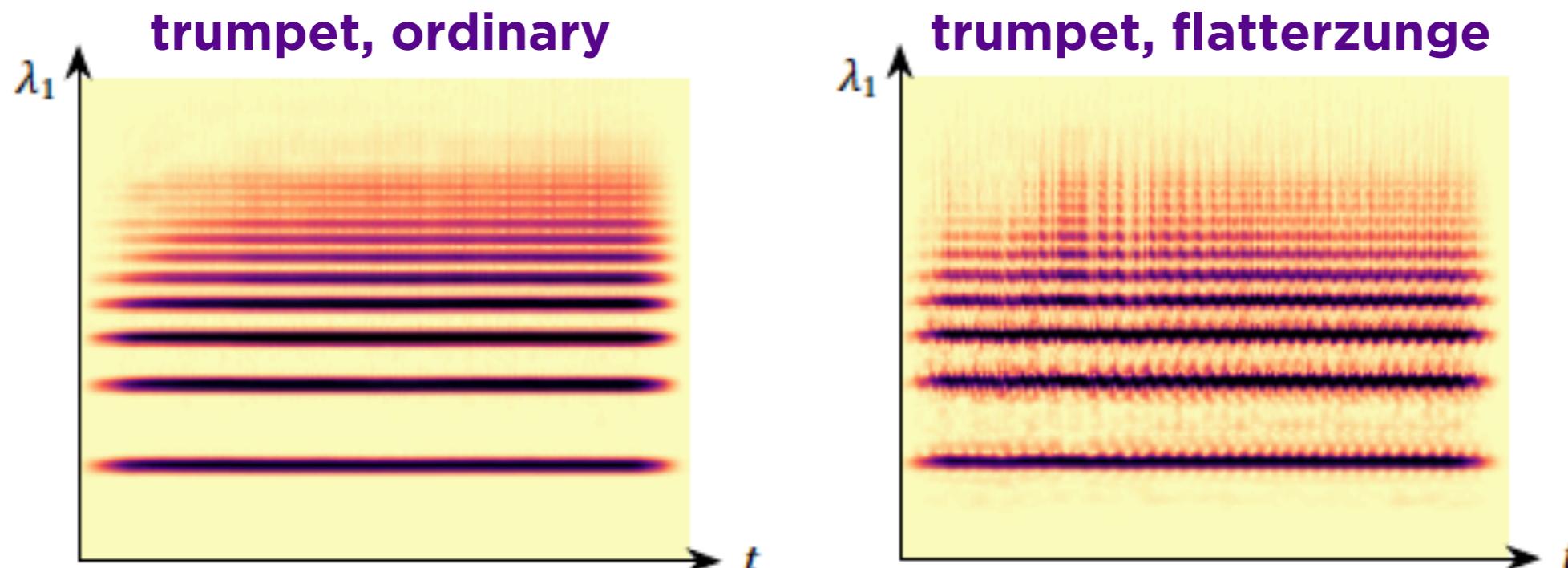
Some, but not all, of these labels are shared with other instruments.



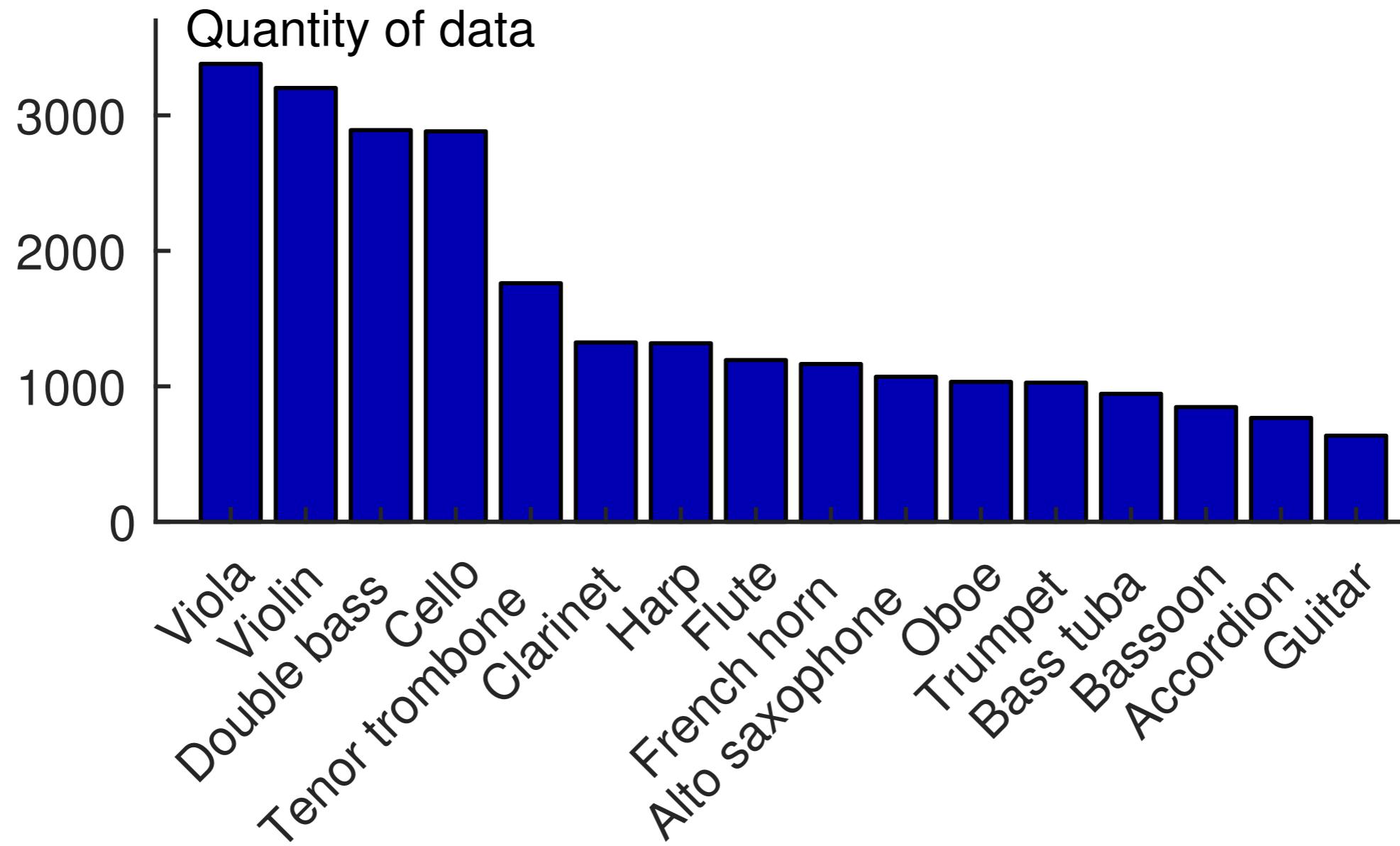
# Why techniques are interesting

1. **Nonstationarity** (transientness):  
spectral envelope changes through time,  
hence a shortcoming of MFCC features.

2. **Long-range interactions** (*Gestalt*):  
variability is not elastic in spectrogram domain,  
hence a shortcoming of Euclidean distances.



# Studio On Line dataset (SOL)



[Ircam 2000]

ordinario  
non-vibrato  
tremolo  
flatterzunge  
sforzando  
crescendo  
note-lasting  
pizzicato-l-vib  
glissando  
decrescendo  
pizzicato-secco  
staccato  
crescendo-to-decrescendo  
ordinario-1q  
trill-minor-second-up  
trill-major-second-up  
sul-ponticello  
pizzicato-bartok  
sul-tasto  
sul-ponticello-tremolo  
multiphonics  
sul-tasto-tremolo  
col-legno-battuto  
col-legno-tratto  
harmonic-fingering  
bisbigliando  
lip-glissando  
artificial-harmonic  
ordinario-to-flatterzunge  
artificial-harmonic-tremolo  
flatterzunge-to-ordinario  
vibrato  
crushed-to-ordinario  
ordinario-to-sul-tasto  
ordinario-to-crushed  
slap-pitched  
sul-ponticello-to-ordinario  
sul-ponticello-to-sul-tasto  
ordinario-to-tremolo  
sul-tasto-to-ordinario  
tremolo-to-ordinario  
near-the-board  
sul-tasto-to-sul-ponticello  
ordinario-to-sul-ponticello  
aeolian-and-ordinario  
brassy  
backwards  
ordinario-high-register  
brassy-to-ordinario  
natural-harmonics-glissandi

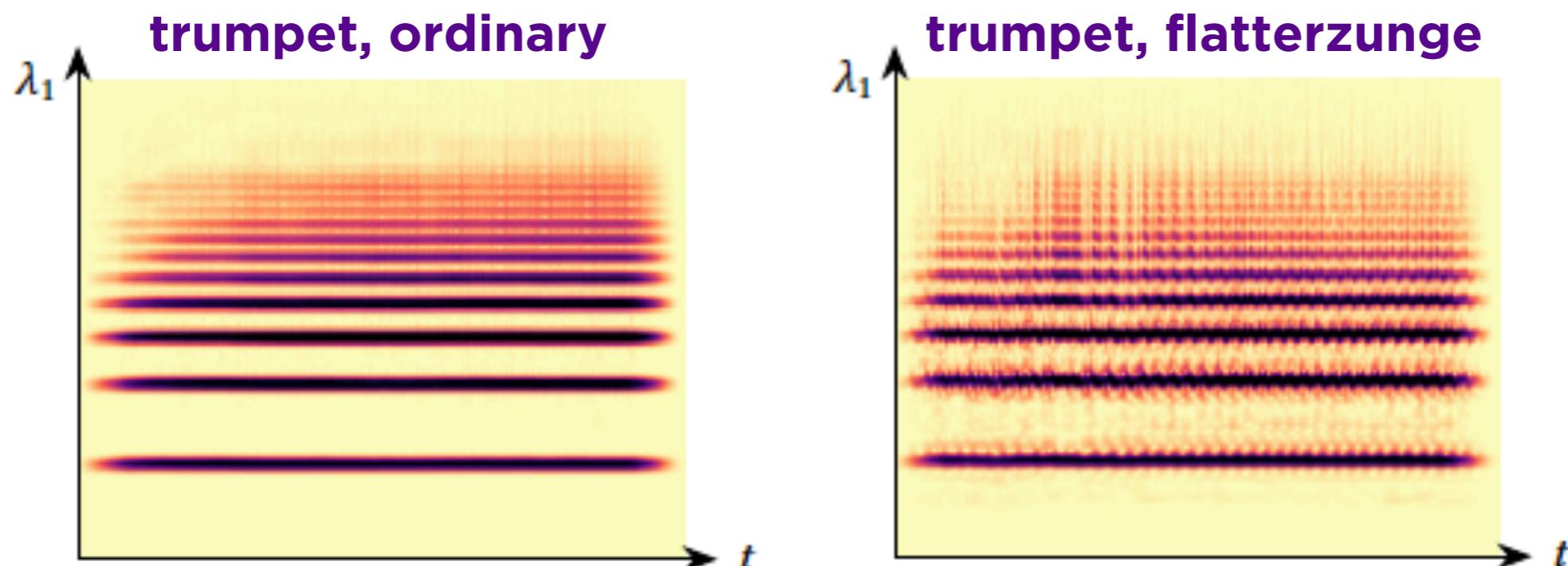
ordinario-to-brassy  
discolored-fingering  
near-the-pegs  
pizzicato  
damped  
near-the-board-with-nail  
key-click  
move-bell-from-down-to-up  
move-bell-from-left-to-right  
glissando-with-stick  
bisbigliando-with-stick  
xylophonic  
open-to-closed  
ordinario-closed  
ordinario-open  
closed-to-open  
combination-of-registers  
dedillo  
flatterzunge-open  
pedal-tone  
combination-of-registers-dbl  
play-and-sing-unison  
slide  
open-to-stopped  
aeolian  
stopped-to-open  
flatterzunge-stopped  
glissando-with-throat  
stopped  
harmonic-in-wood  
aeolian-to-ordinario  
ordinario-to-aeolian  
single-tonguing  
flatterzunge-and-voice-unison  
double-tonguing  
flatterzunge-closed  
play-and-sing-m2-up  
half-valve-glissando  
blow-without-reed  
whistle-tones-sweeping  
double-trill-minor-second-up  
glissando-with-pedal  
slap-unpitched  
play-and-sing  
play-and-sing-M7-up  
play-and-sing-P5-up  
play-and-sing-aug4-up  
behind-the-bridge  
hit-on-body  
buzz

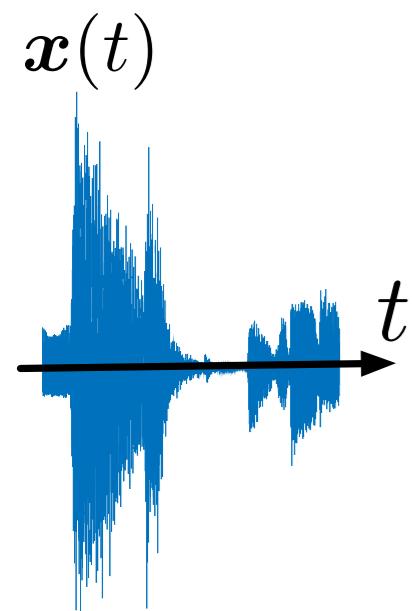
scratch-with-nail  
tap-with-stick  
on-the-tuning-pegs  
tongue-ram  
buzzing-pedal  
exploding-slap-pitched  
double-trill-major-second-up  
flatterzunge-high-register  
harmonics-glissando  
kiss  
ordinario-no-mouthpiece  
whistle-tones  
cluster  
discolored-fingering-1q  
tap-on-body  
speak-into-instrument  
bottleneck  
cluster-with-nail  
glissando-near-the-board  
glissando-with-nail  
behind-the-fingerboard  
double-glissando  
chromatic-scale  
flatterzunge-no-mouthpiece  
glissando-fluido-with-stick  
on-the-bridge  
on-the-tailpiece  
play-and-sing-glissando  
growl  
on-the-frog  
vocalize-on-harmonics  
behind-the-frog  
bellowshake  
blow  
breath  
drum-effect  
exploding-slap-unpitched  
filtered-by-voice  
glissando-embouchure  
inhaled  
jet-whistle  
percussion-embouchure  
tremolo-with-fingertips

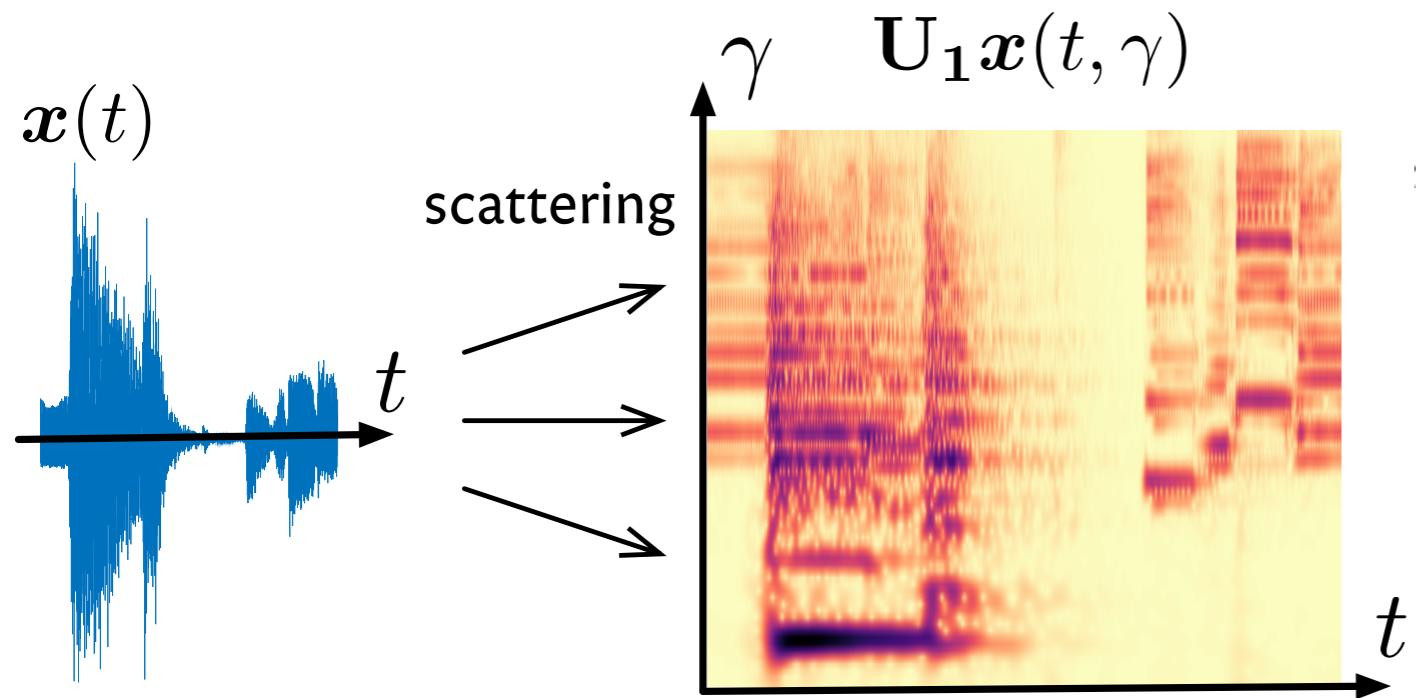
# Contributions

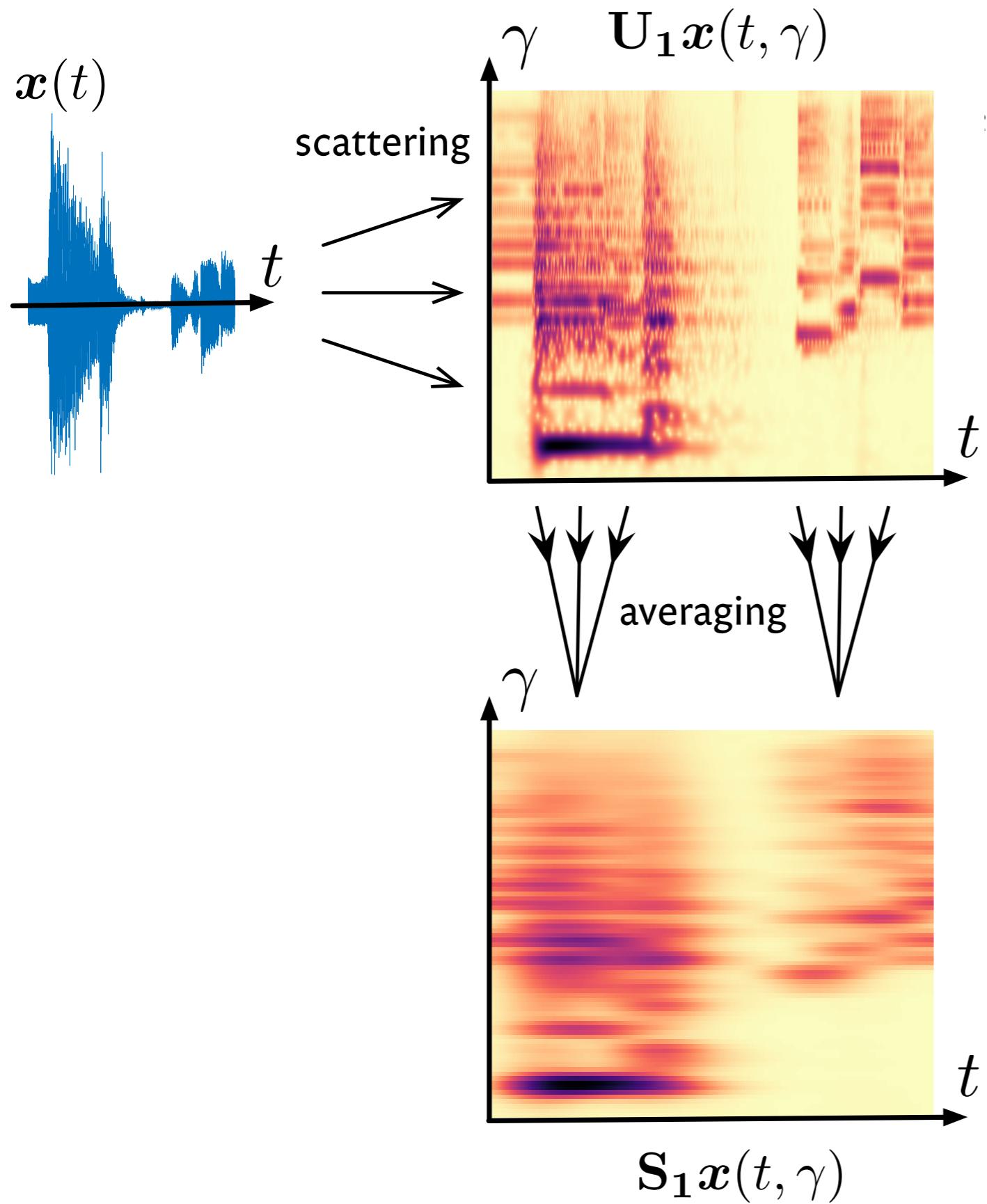
1. **Nonstationarity** (transientness):  
replace MFCC by scattering transform.

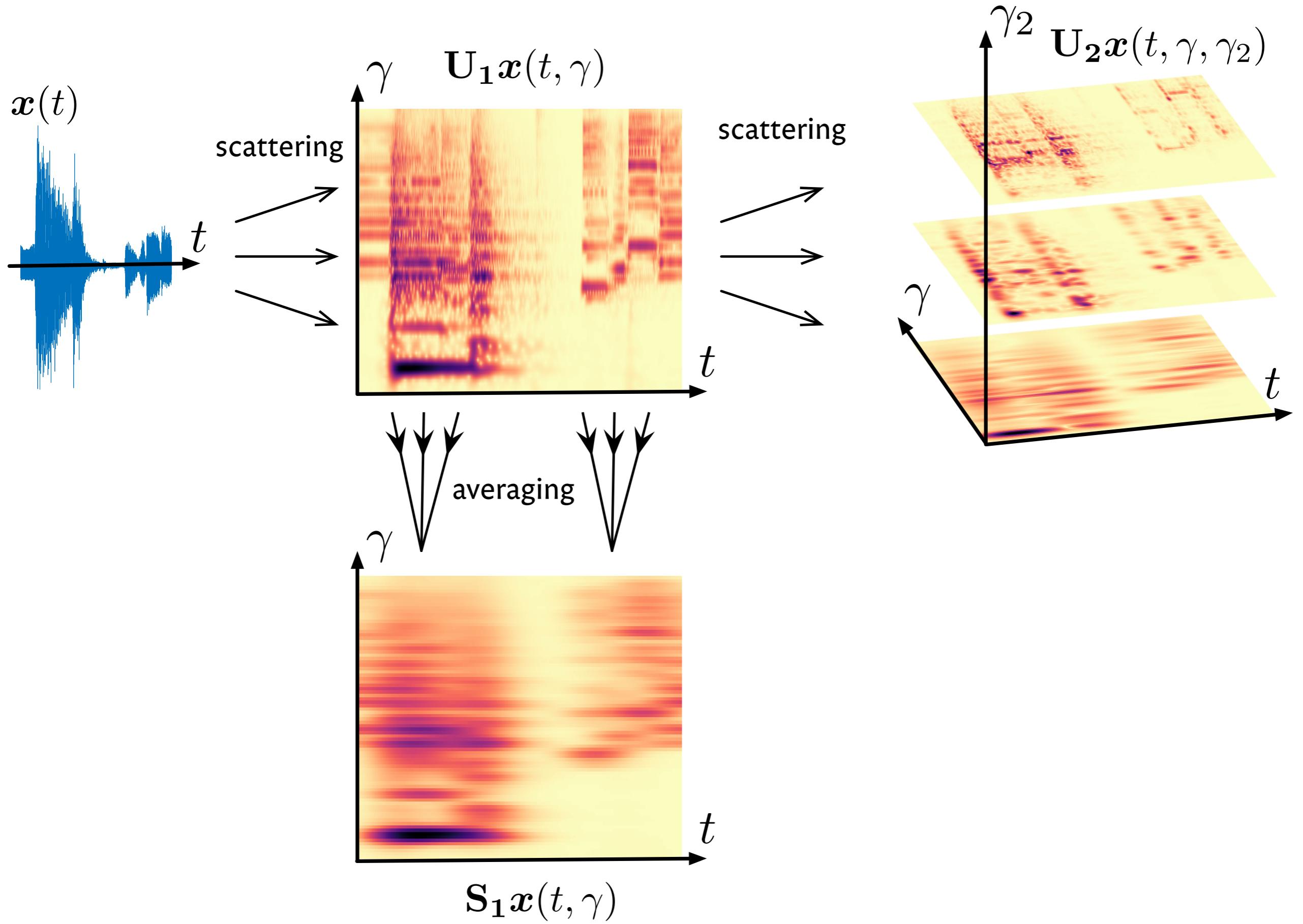
2. **Long-range interactions** (*Gestalt*):  
replace Euclidean neighbors by metric learning.

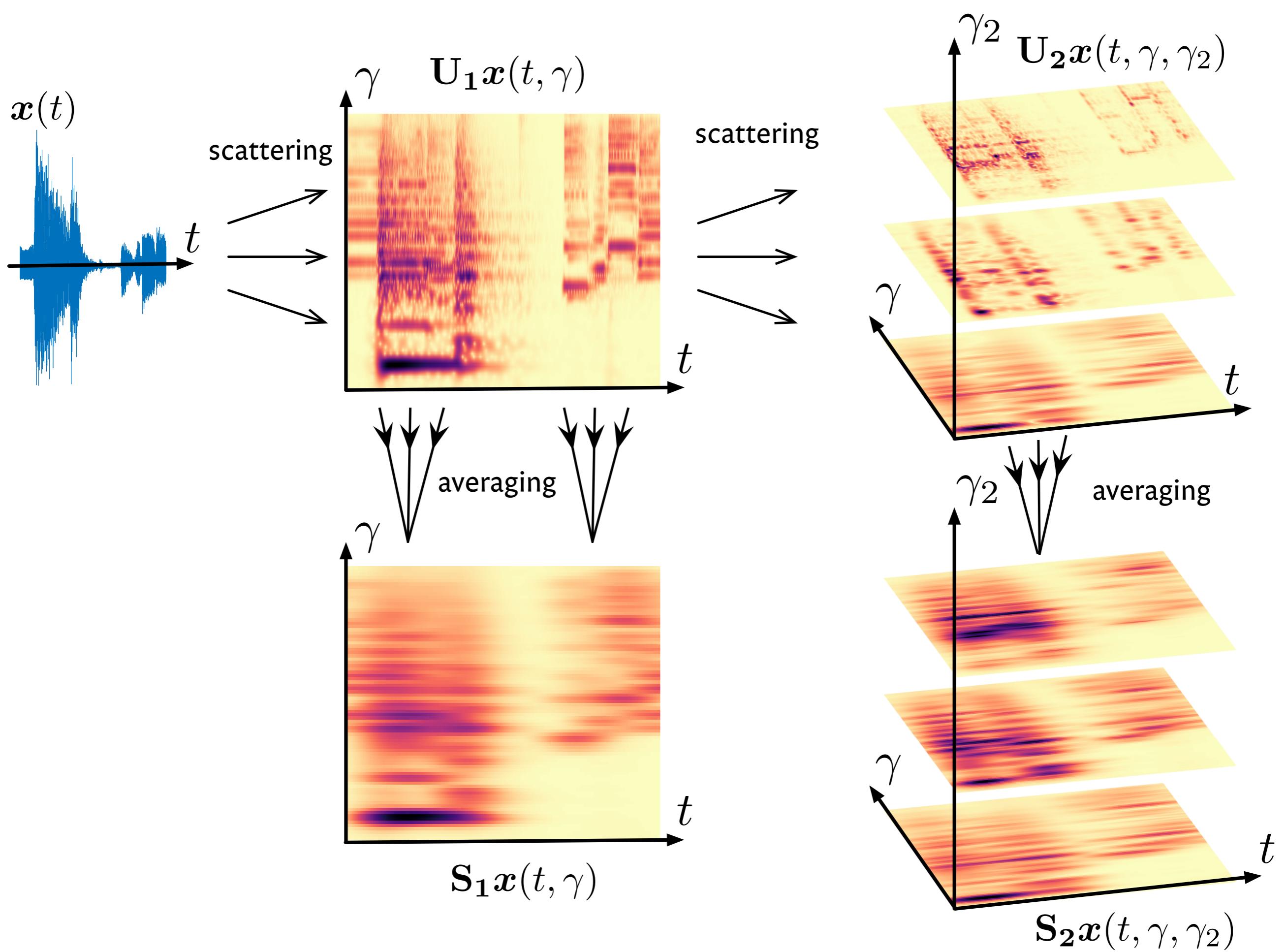










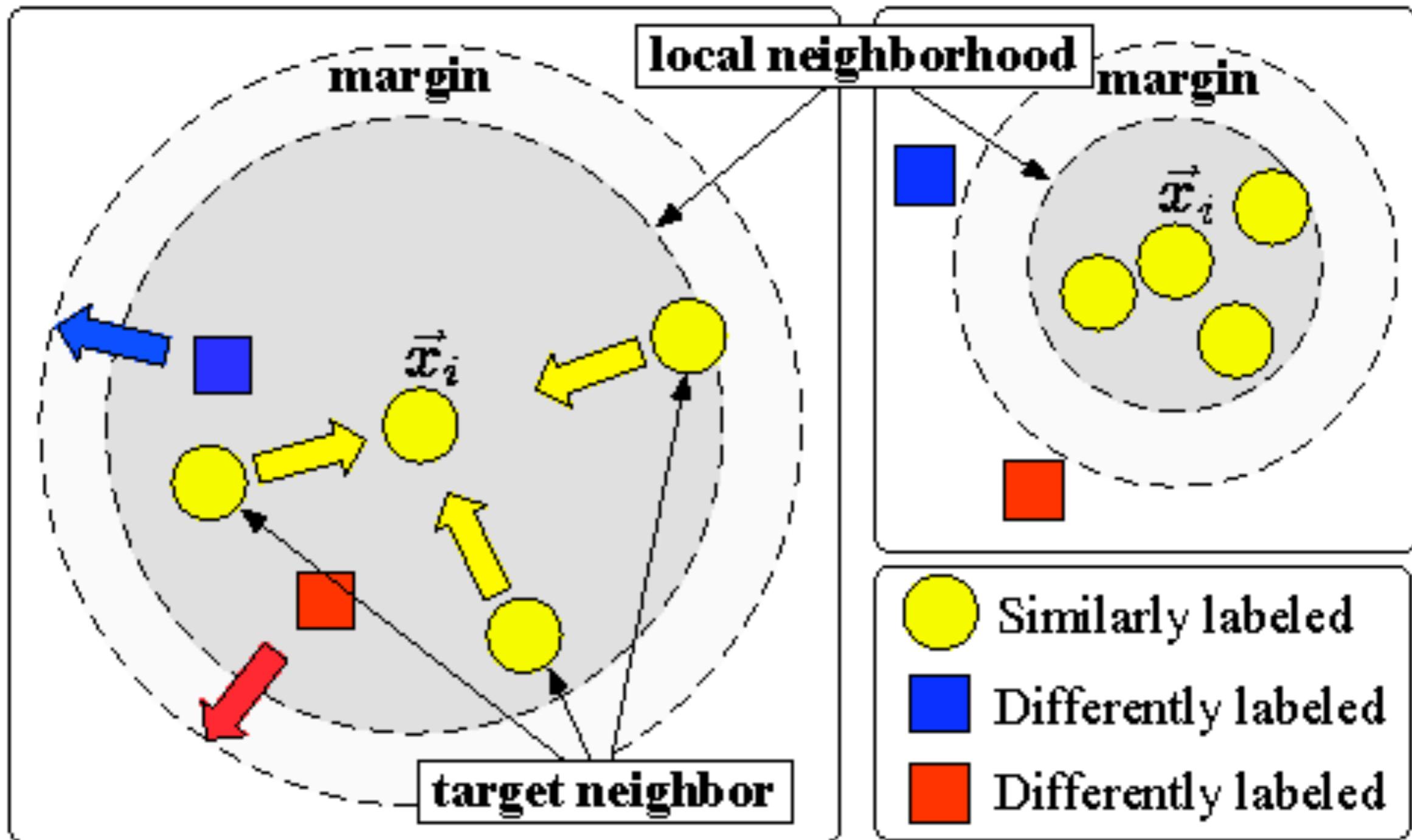


# Large-margin nearest neighbors

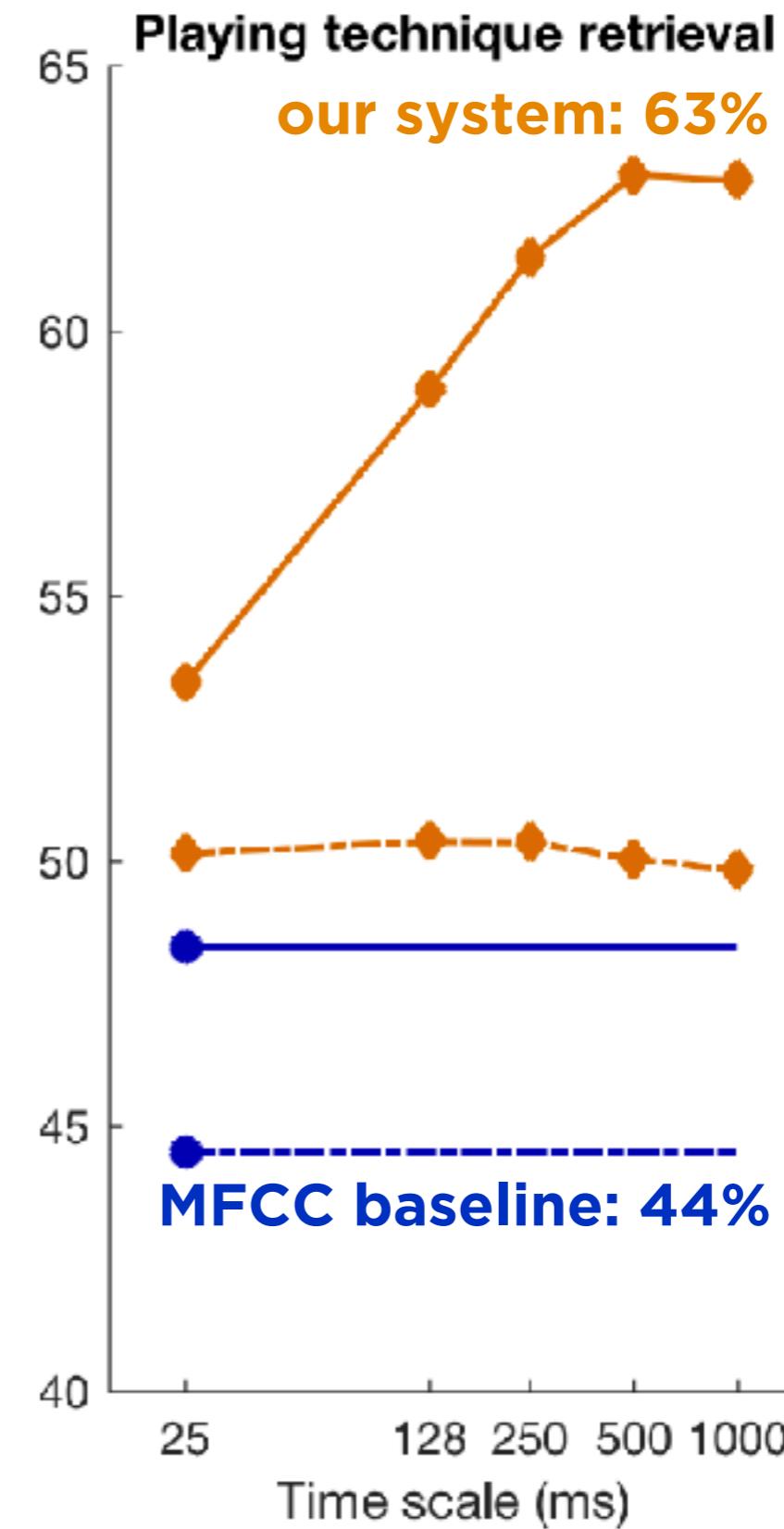
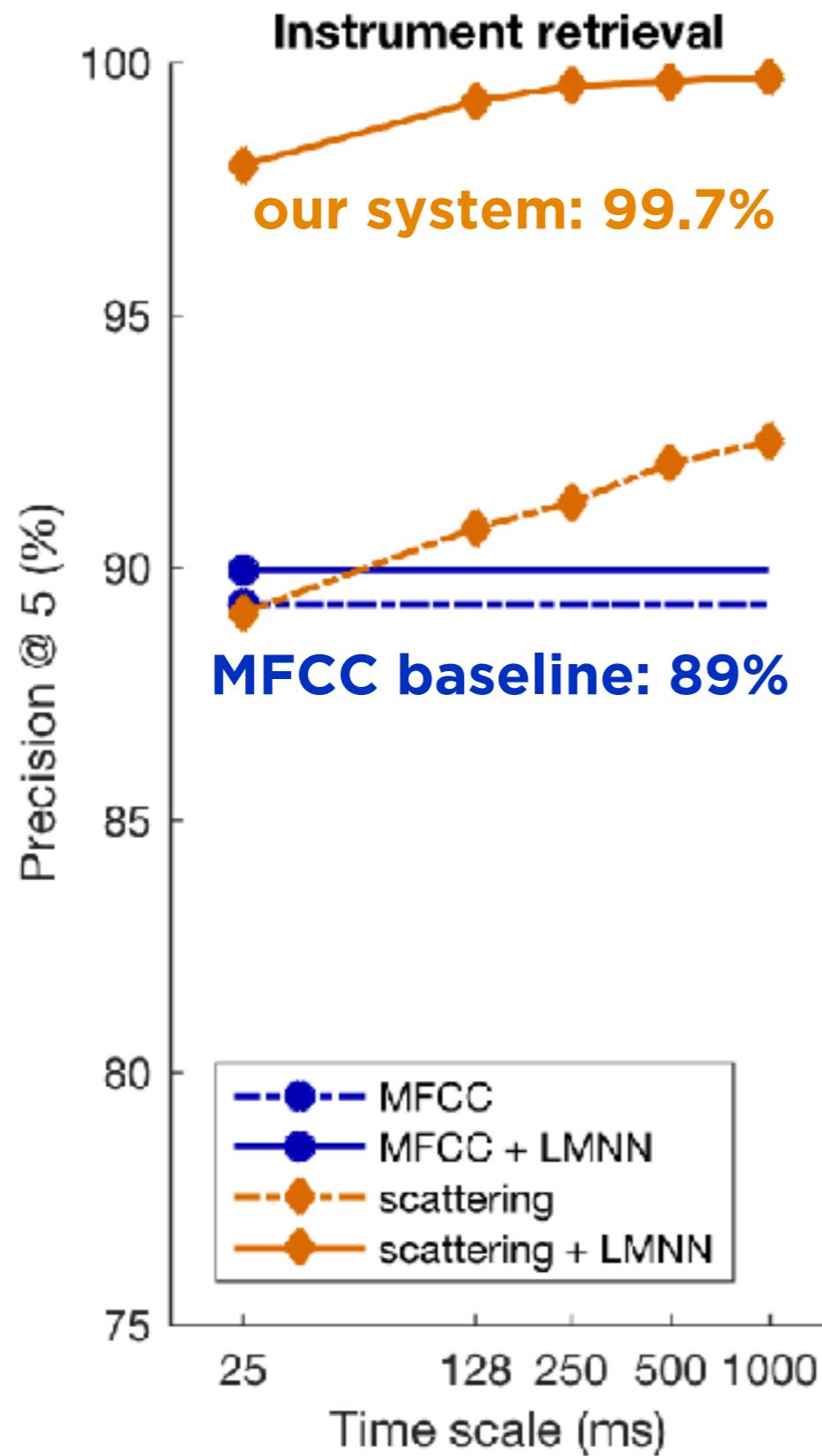
[Weinberger]

BEFORE

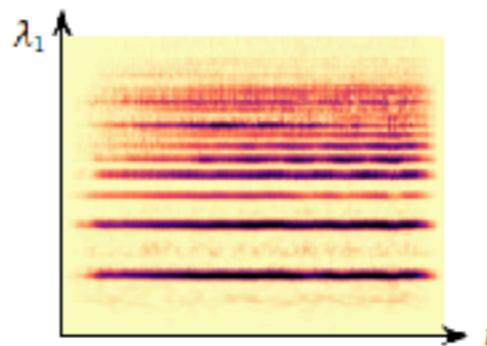
AFTER



# Query-by-example results

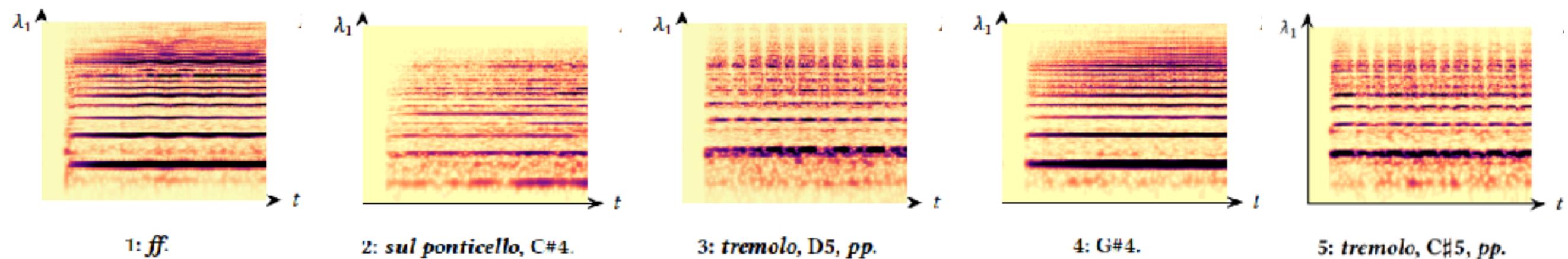


# Demo: case of an ordinary query

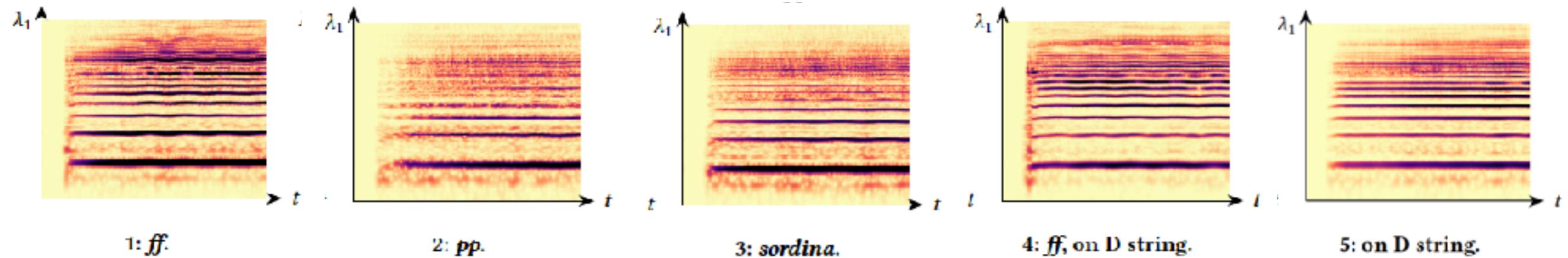


Query: Violin, *ordinario*, G4, *mf*, on G string.

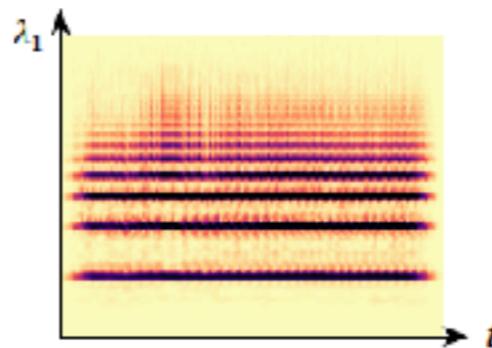
## MFCC neighbors



## scattering transform neighbors

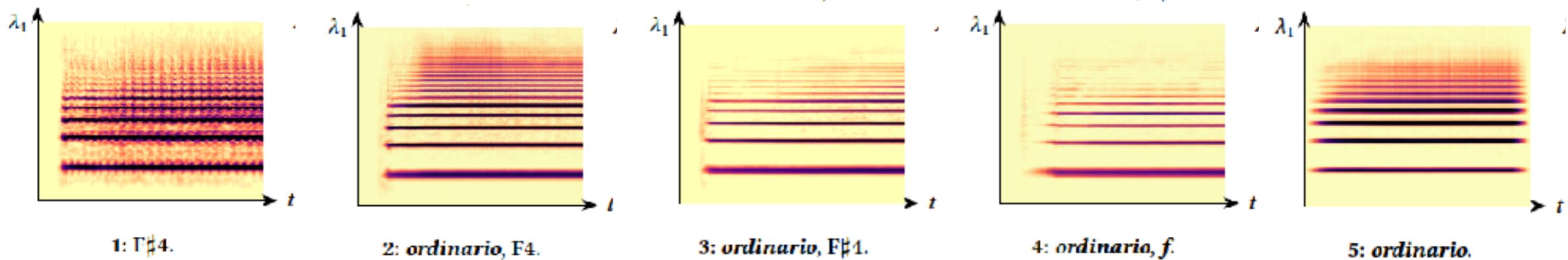


# Demo: case of a flatterzunge query

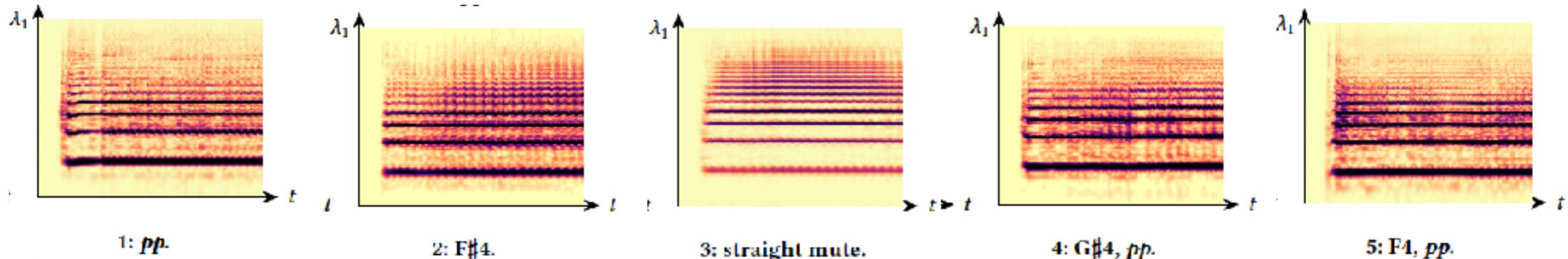


Query: Trumpet in C, *flatterzunge*, G4, *mf*.

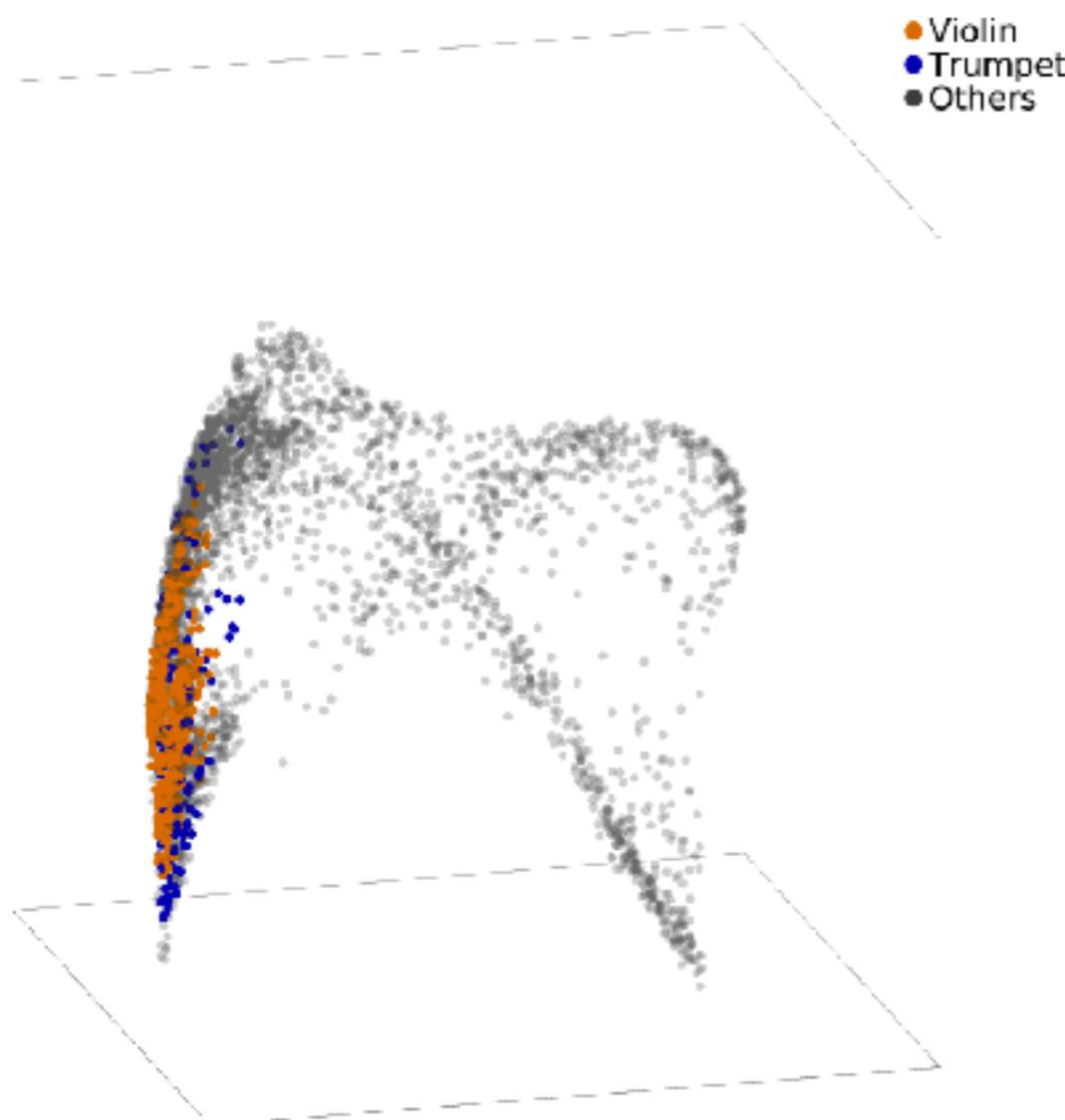
## MFCC neighbors



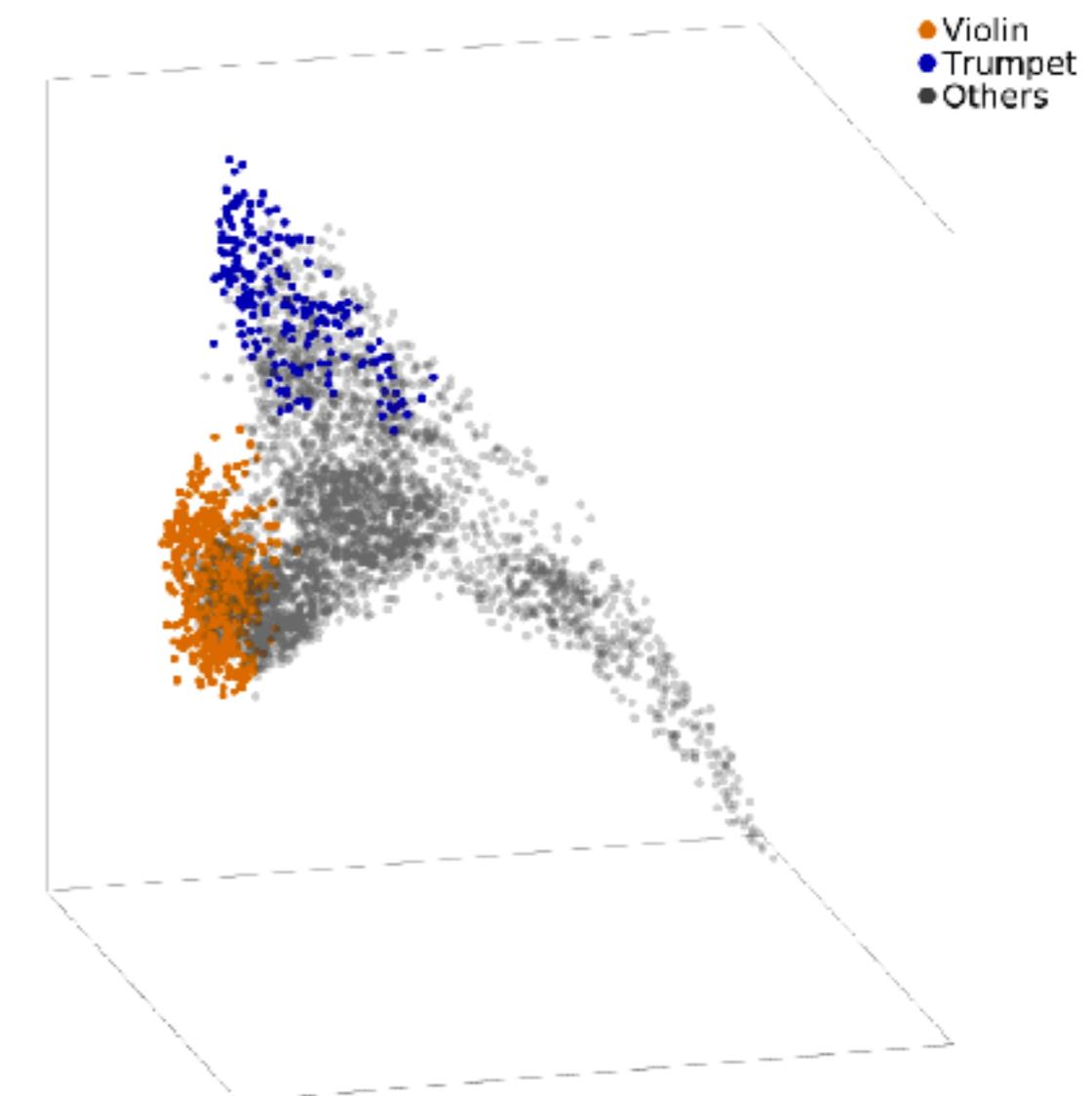
## scattering transform neighbors



# Isomap embedding

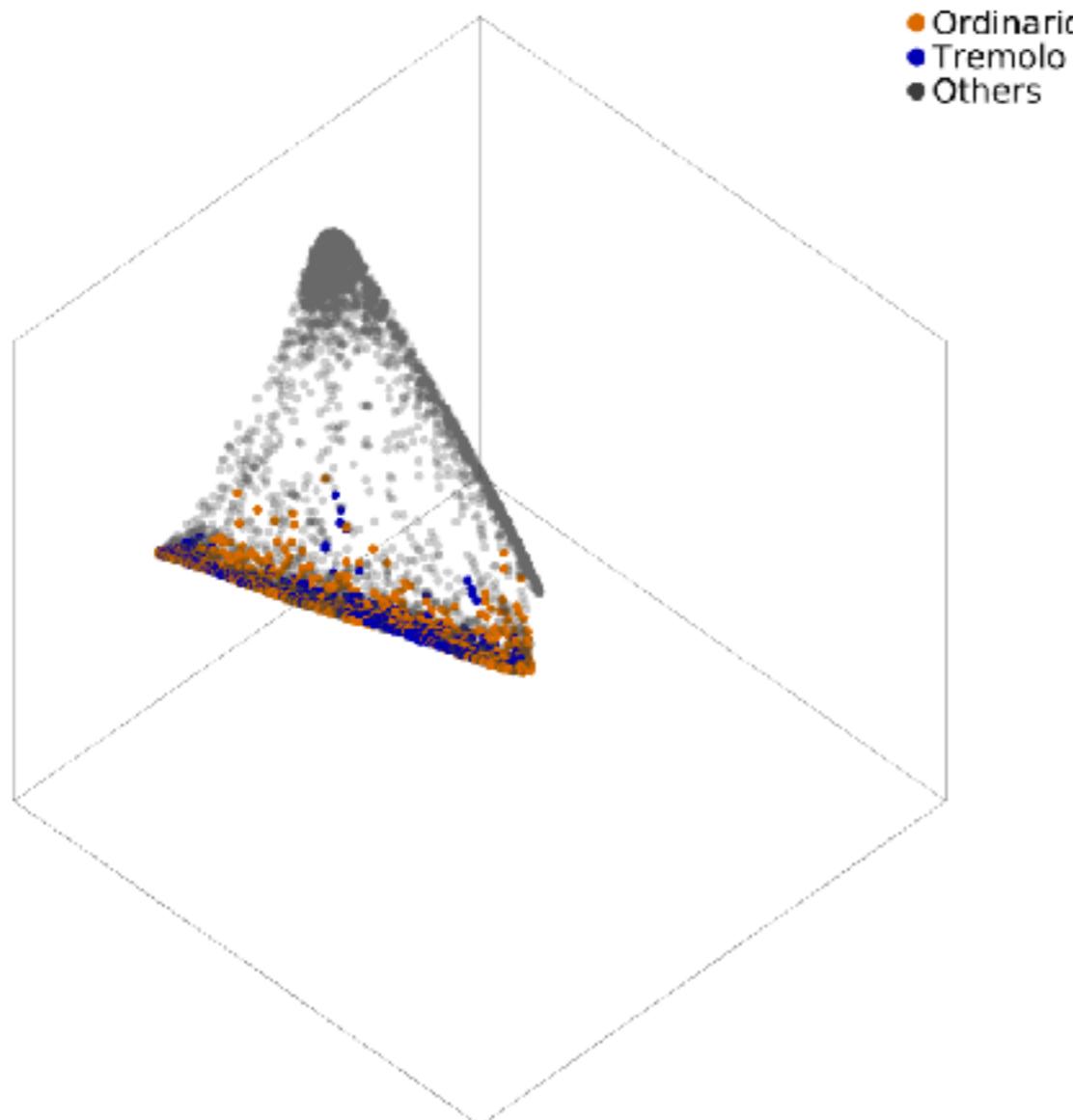


(a) Instrument embedding with MFCC.

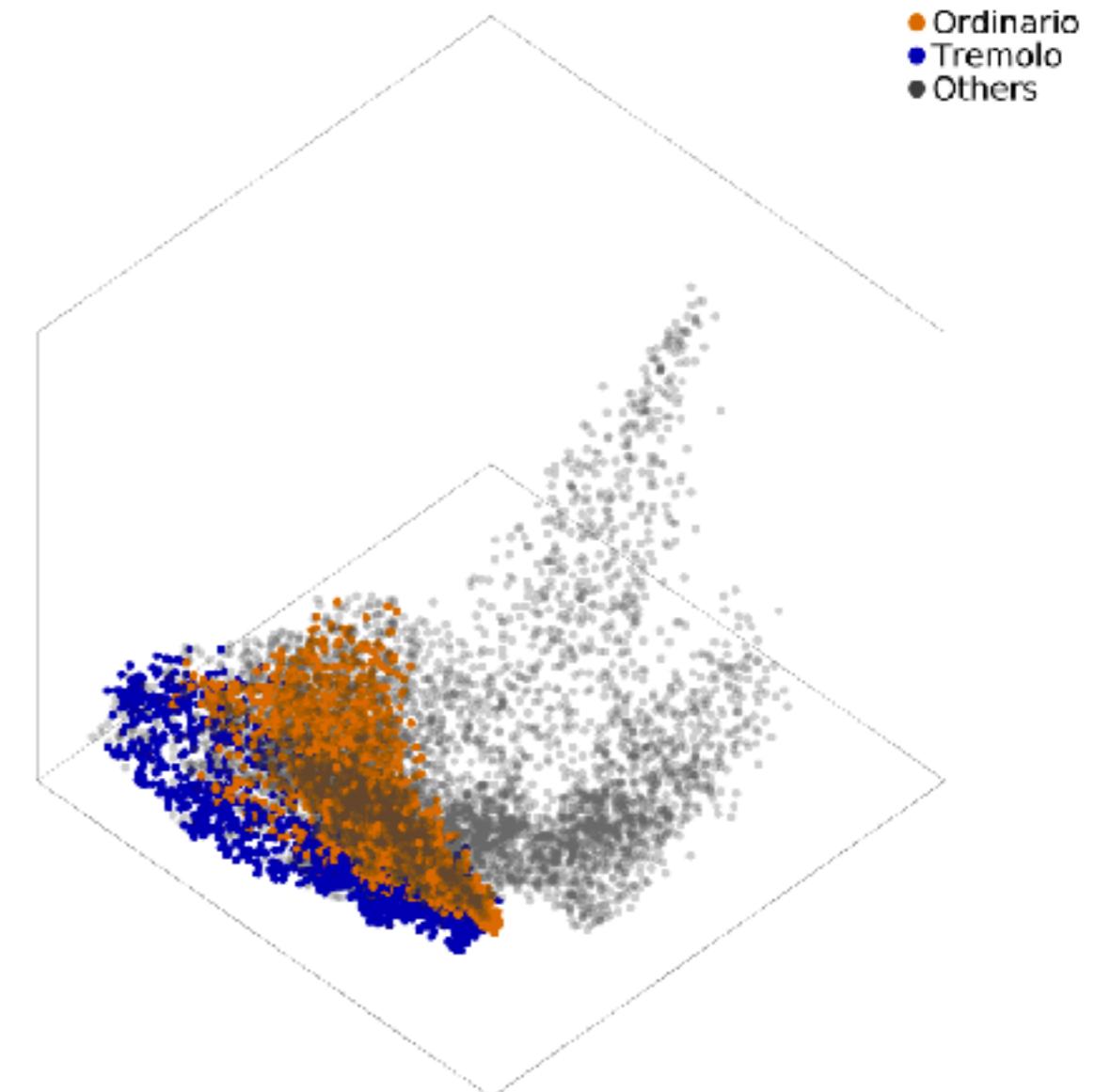


(b) Instrument embedding with scattering transform.

# Isomap embedding



(c) Playing technique embedding with MFCC.



(d) Playing technique embedding with scattering transform.

Take-home message:  
the timbral modeling of musical events  
is **only “solved” for ordinary techniques**,  
but is still challenging  
for extended techniques

*L'effet est plus important que les notes* – Levinas

