Universiteit Utrecht

NWO
Netherlands Organisation for Scientific Research

Melodic Segmentation Across Cultural Traditions

MARCELO E. RODRÍGUEZ-LÓPEZ & ANJA VOLK

Department of Information and Computing Sciences
Utrecht University, Interaction Technology Group
M.E.RodriguezLopez@uu.nl

June 13, 2014

Universiteit Utrecht

NWO
Netherlands Organisation for Scientific Research

What is Segmentation

What is Segmentation

$$\text{Segmenting} = \text{Chunking} = \text{Grouping}$$

What is Segmentation

$$\text{Segmenting} = \text{Chunking} = \text{Grouping}$$

Psychology:
process of **grouping** individual units of information into **larger** units.

What is Segmentation

$$\text{Segmenting} = \text{Chunking} = \text{Grouping}$$

Psychology:
process of **grouping** individual units of information into **larger** units.

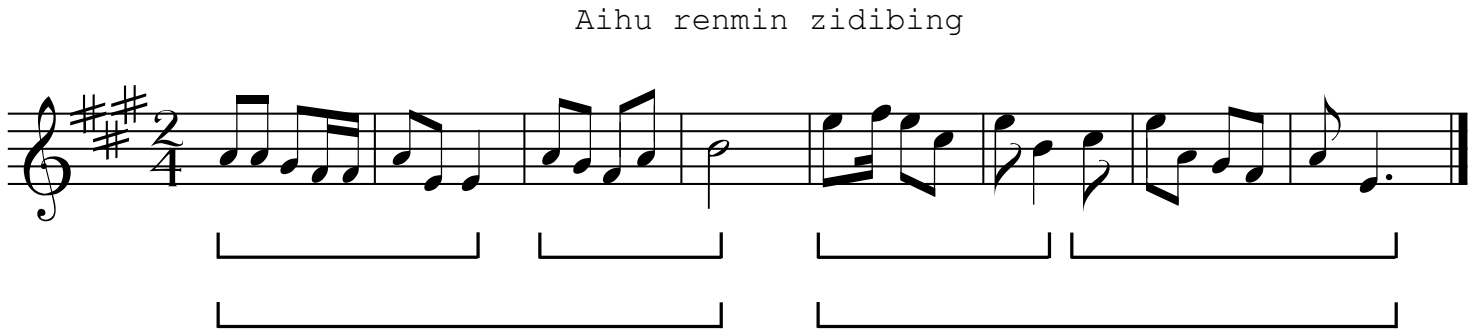
Computing:
process of **dividing** data into **smaller** meaningful units.

Music Segmentation: DM Task Definition

In **Digital Musicology (DM)**, segmentation is the task of **dividing** a musical piece/melody/section into 'smaller' **structural units**.

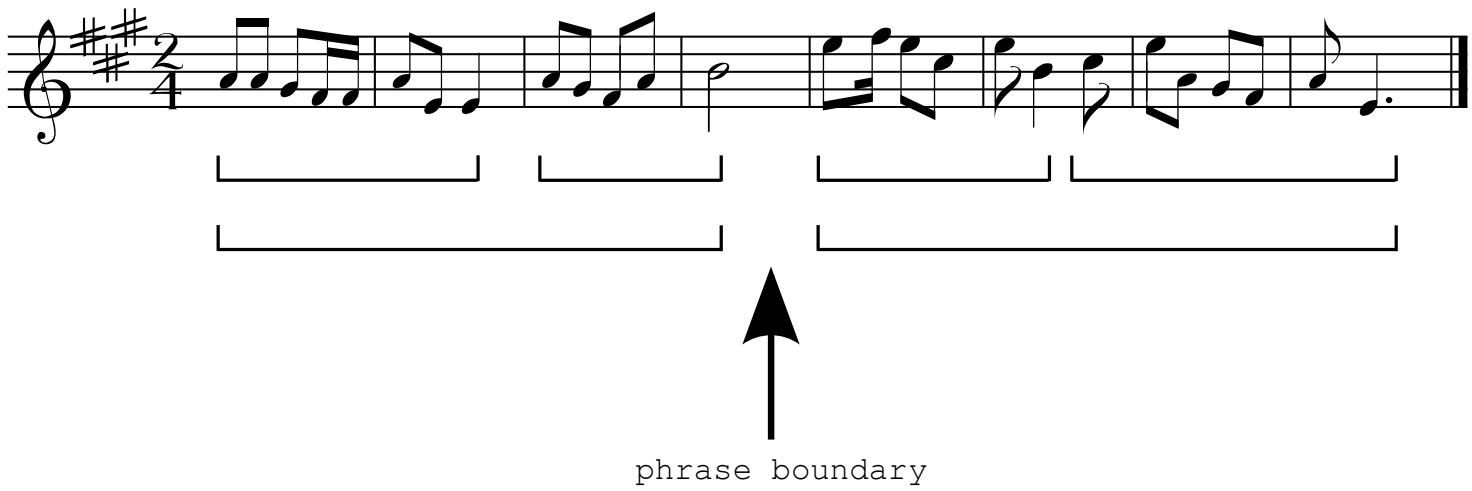
Music Segmentation: Scope

- **Type of Music and Segment Granularity**
  - **input**: monophonic music (most generally melodies)
  - **encoding**: symbolic
  - **music-theoretic segment parallel**: phrases

Melody Segmentation: an Example

Melody Segmentation: Task

- **Task**
  - **identify** segment boundary locations
  - ~~**pair** boundaries (begin, end)~~
  - ~~**label** boundaries~~

Melody Segmentation: Task



Models of Melody Segmentation

- **development**: +30 years
- **# of models**: ≈27
- **comparative studies**: 4, 3-6 models evaluated
- **most successful**: Gestalt-based

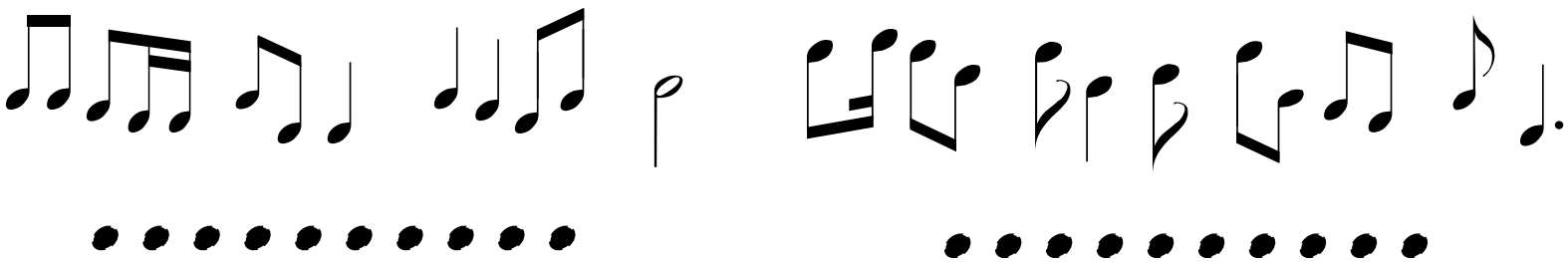
Gestalt Models: Basic Overview

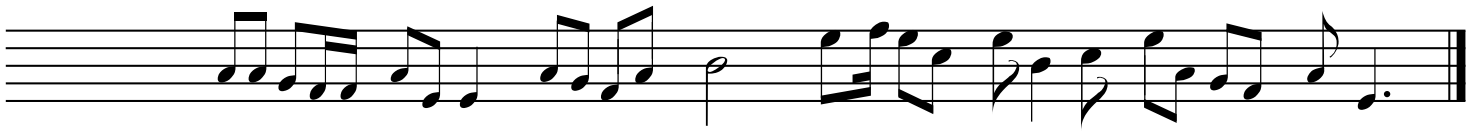
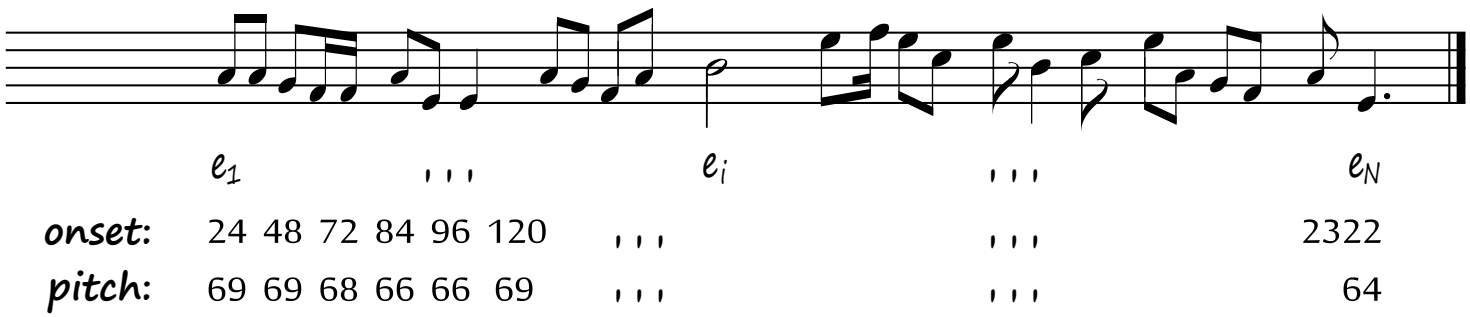
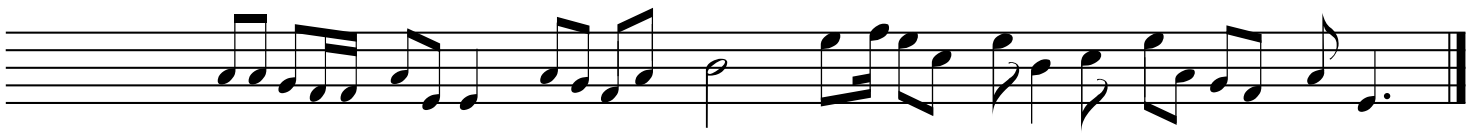
- attempt to quantify (visual) Gestalt principles
- use system of preference rules
- **Gestalt proximity is modeled as 'discontinuity' detection**

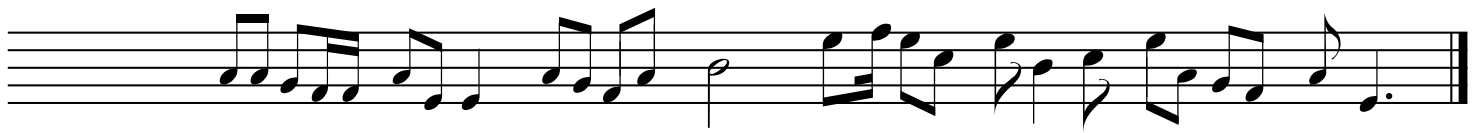
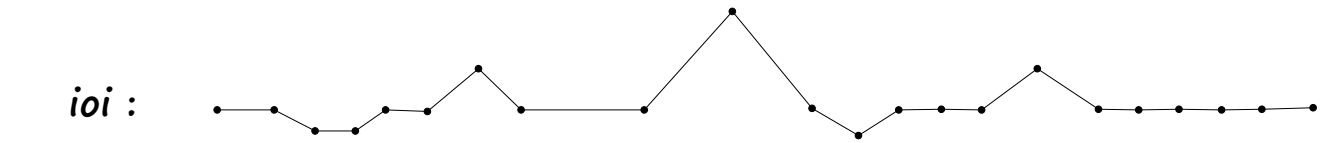
(breaks in the melodic flow)

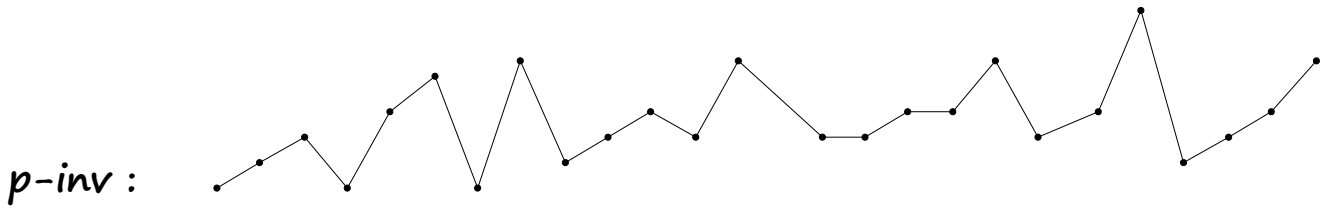
Local Discontinuities Detection: an Example

Local Discontinuities Detection: an Example

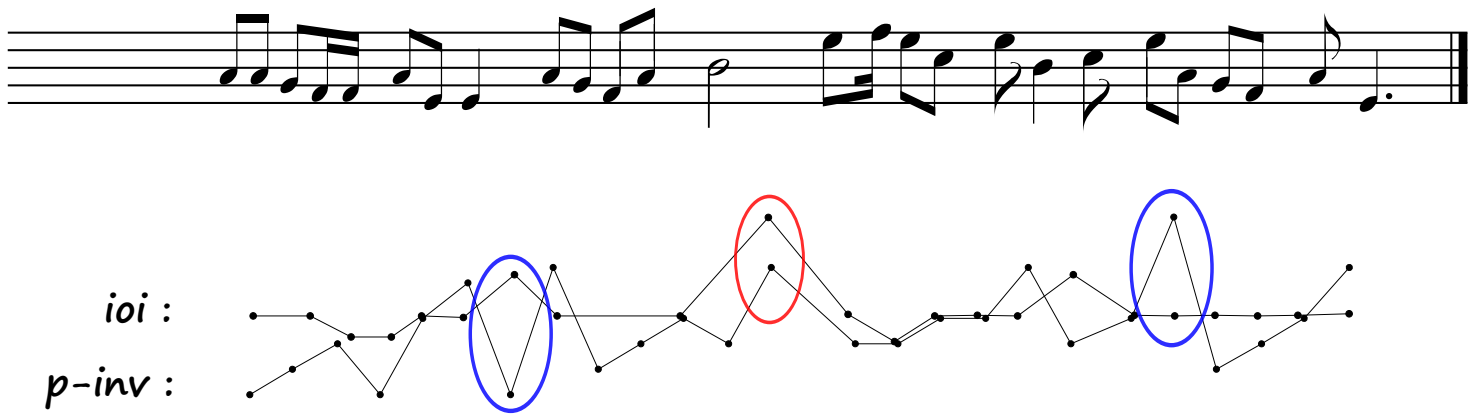


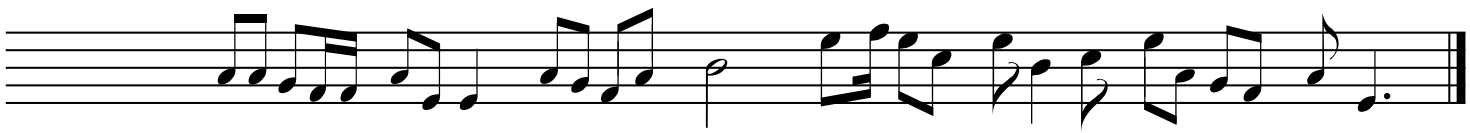
Local Discontinuities Detection: an Example

Local Discontinuities Detection: Input

Local Discontinuities Detection: Input

ioi :	.5	.5	.5	.25	.25	.5	.5	1	.5	.5	.5	.5	2	.75	.25	.5	.5	.5	1	.5	.5	.5	.5	.5	.5
p-inv :	0	1	2	0	3	5	0	5	1	2	3	2	5	2	2	3	3	5	2	3	7	1	2	3	5

Local Discontinuity Detection: Profiles





Local Discontinuity Detection: Profiles



Local Discontinuity Detection: Output

Output : 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

Gestalt Models: Assumptions

- discontinuity as a **distance** in some perceptual space
- discontinuity can be treated as a **local phenomenon**
- discontinuity is **universal/idiom-independent**

Statistical Study: Research Questions

- relevance of discontinuity as a boundary cue?
- relevance of discontinuity parametric representation?
- appropriateness of locality for modeling?
- appropriateness of universality assumption?

Statistical Study: Data

ESSEN FOLK SONG COLLECTION (EFSC)

- 20,000 annotated vocal folk melodies
- Representative subsets of German and Chinese songs

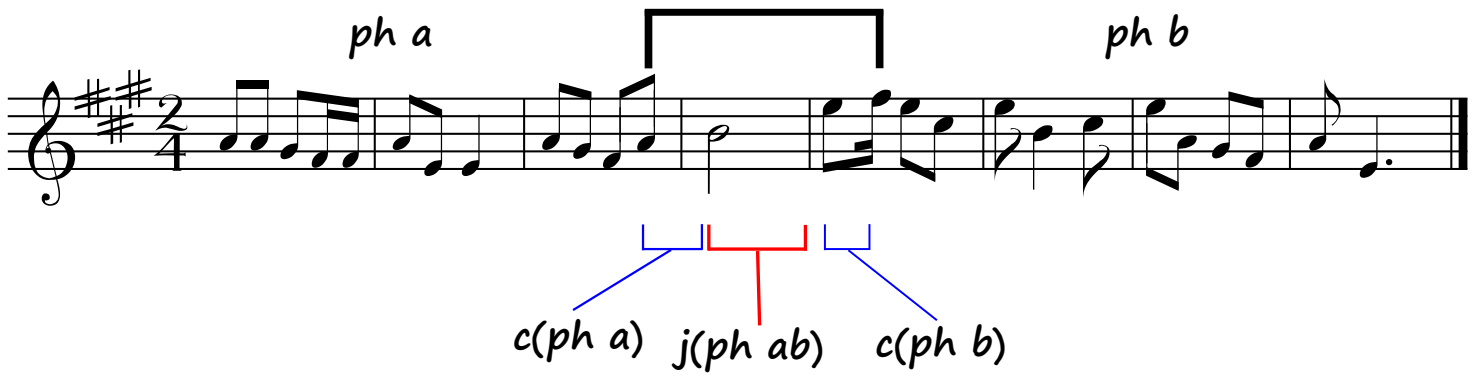
EXPERIMENTAL SUBSETS

- $\approx 5k$ German songs, $\approx 30k$ phrases, $1 \ldots 29$ notes
- $\approx 2k$ Chinese songs, $\approx 12k$ phrases, $1 \ldots 169$ notes

FILTERED

- $\approx 1.2k$ German songs, $\approx 5.5k$ phr-pairs, $13 \ldots 21$ notes
- $\approx 1.4k$ Chinese songs, $\approx 4.6k$ phr-pairs, $2 \ldots 28$ notes

Statistical Study: Processing



Nioi :	.35	.5	.15
p-inv :	2	5	2

Results 1:

pitch	duration	Germany	China
d	-	32.41 %	43.73 %
-	*d*	49.57 %	72.56 %
d	-	04.51 %	22.53 %
-	*d*	24.13 %	57.70 %
nd	-	31.49 %	28.53 %
-	*nd*	23.26 %	09.04 %

Results 1:

pitch	duration	Germany	China
d	-	32.41 %	43.73 %
-	*d*	49.57 %	**72.56** %
d	-	04.51 %	22.53 %
-	*d*	24.13 %	57.70 %
nd	-	31.49 %	28.53 %
-	*nd*	23.26 %	09.04 %

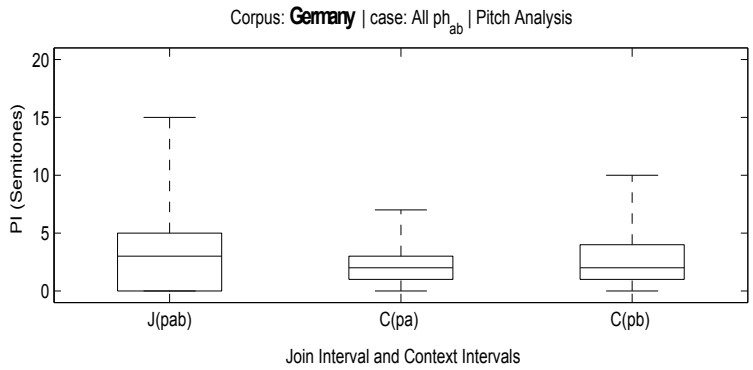
Results 1:

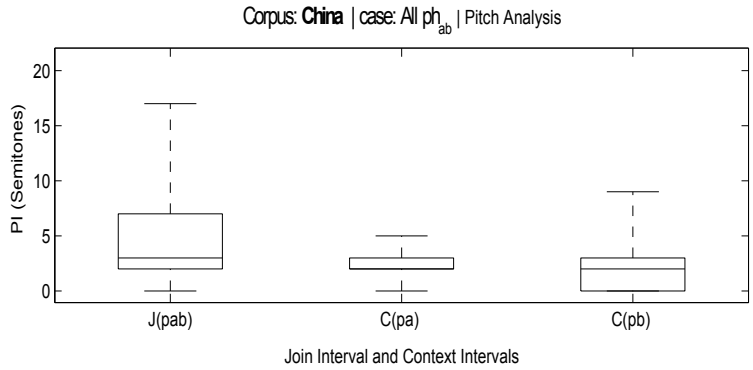
pitch	duration	Germany	China
d	-	32.41 %	43.73 %
-	*d*	49.57 %	72.56 %
d	-	04.51 %	22.53 %
-	*d*	24.13 %	**57.70** %
nd	-	31.49 %	28.53 %
-	*nd*	23.26 %	**09.04** %

Results 1:

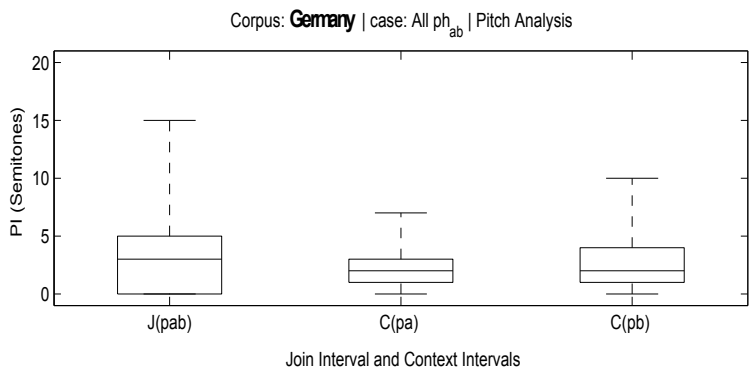
pitch	duration	Germany	China
d	-	32.41 %	43.73 %
-	*d*	49.57 %	72.56 %
d	-	**04.51** %	22.53 %
-	*d*	24.13 %	57.70 %
nd	-	31.49 %	28.53 %
-	*nd*	23.26 %	09.04 %

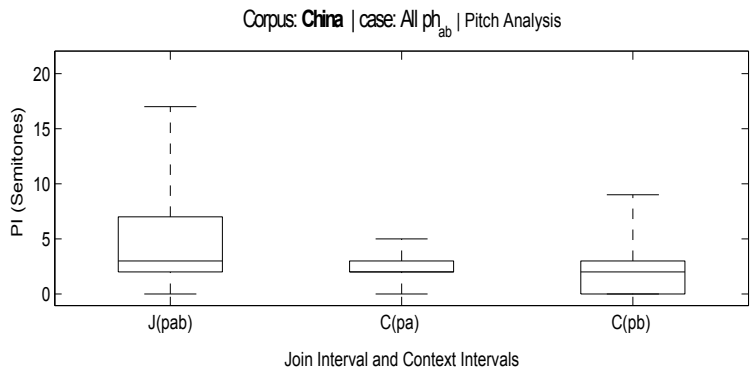
Results 2:

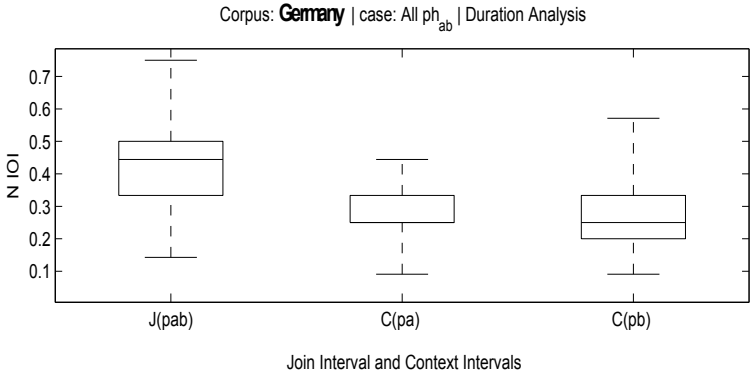

Corpus: Germany | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

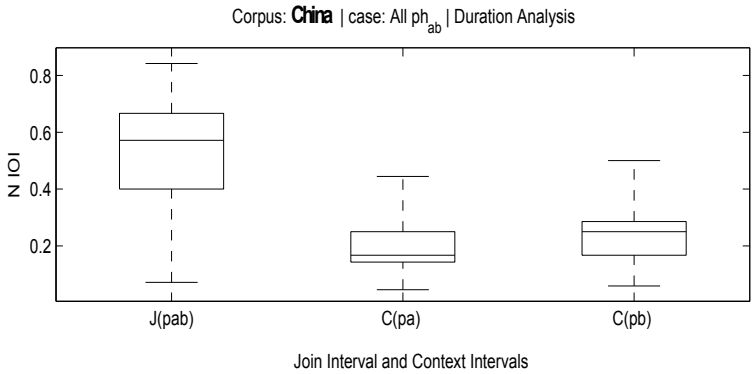

Corpus: China | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

Results 2:

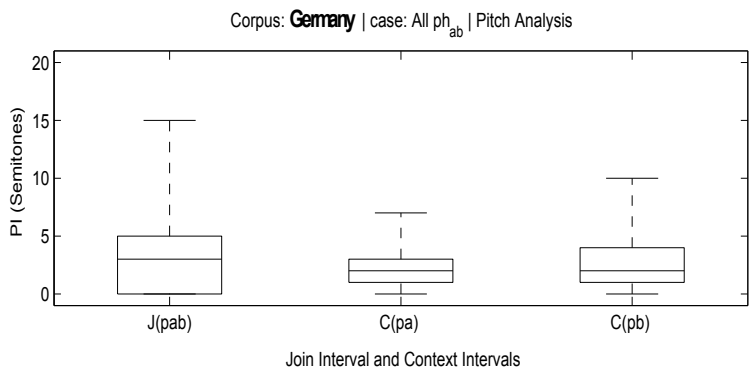

Corpus: Germany | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

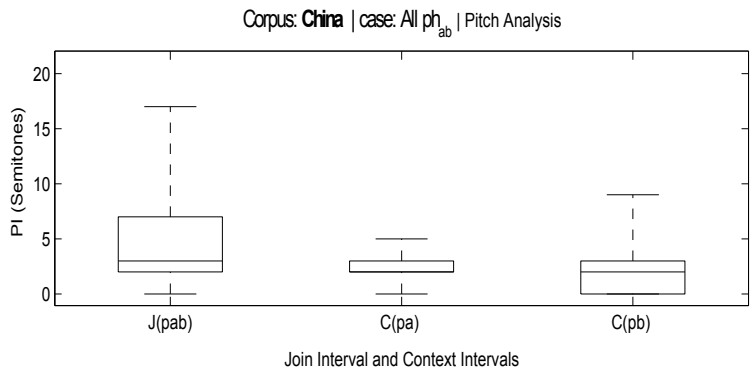

Corpus: China | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

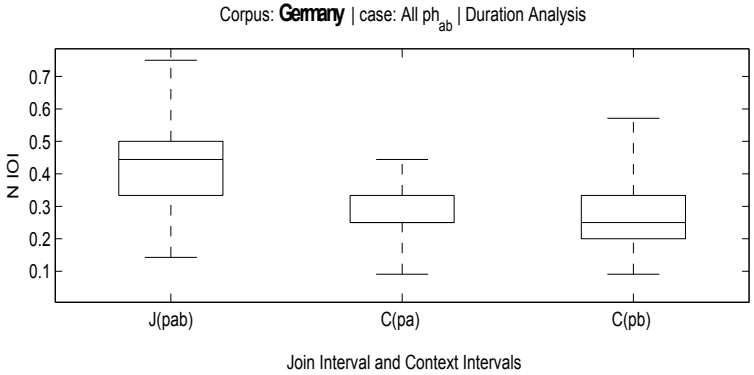

Corpus: Germany | case: All ph_ab | Duration Analysis
N IOI
0.7
0.6
0.5
0.4
0.3
0.2
0.1
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

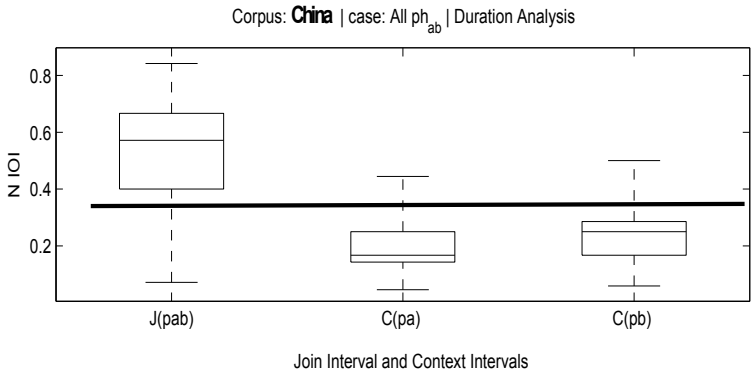

Corpus: China | case: All ph_ab | Duration Analysis
N IOI
0.8
0.6
0.4
0.2
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

Results 2:

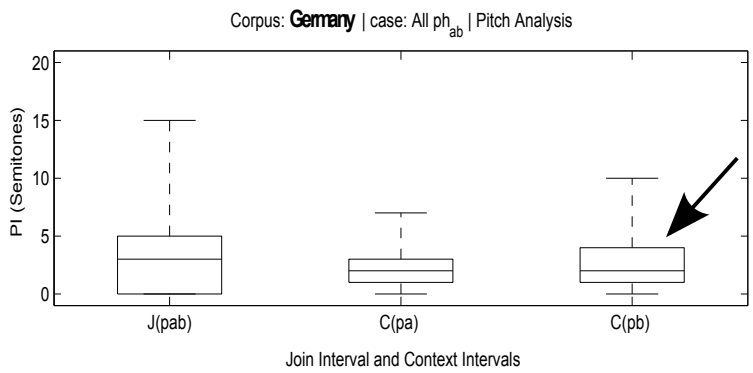

Corpus: Germany | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

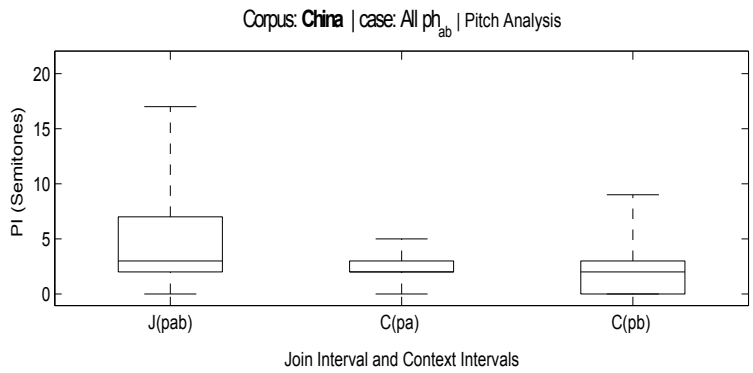

Corpus: China | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

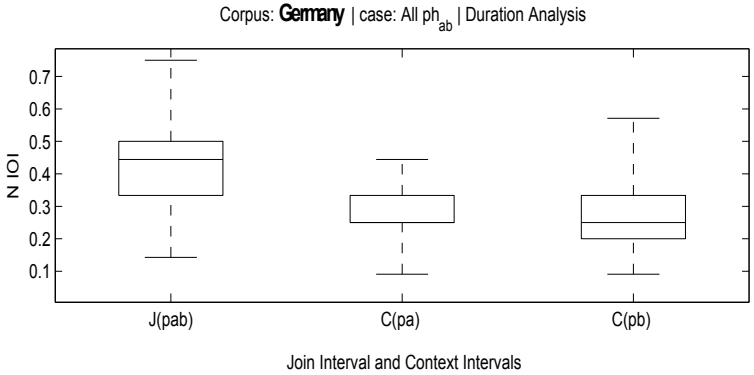

Corpus: Germany | case: All ph_ab | Duration Analysis
N IOI
0.7
0.6
0.5
0.4
0.3
0.2
0.1
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

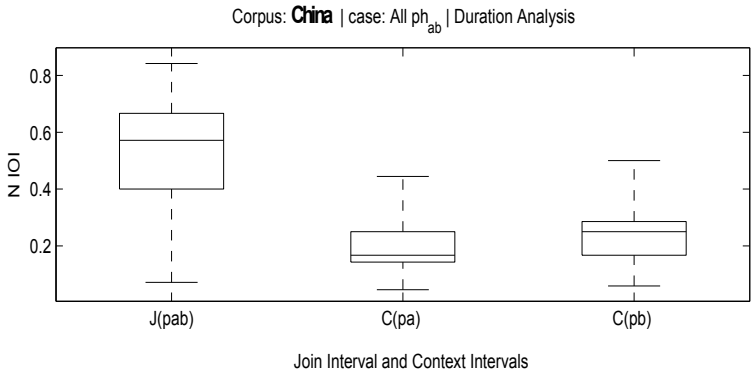

Corpus: China | case: All ph_ab | Duration Analysis
N IOI
0.8
0.6
0.4
0.2
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

Results 2:


Corpus: Germany | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals


Corpus: China | case: All ph_ab | Pitch Analysis
PI (Semitones)
20
15
10
5
0
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

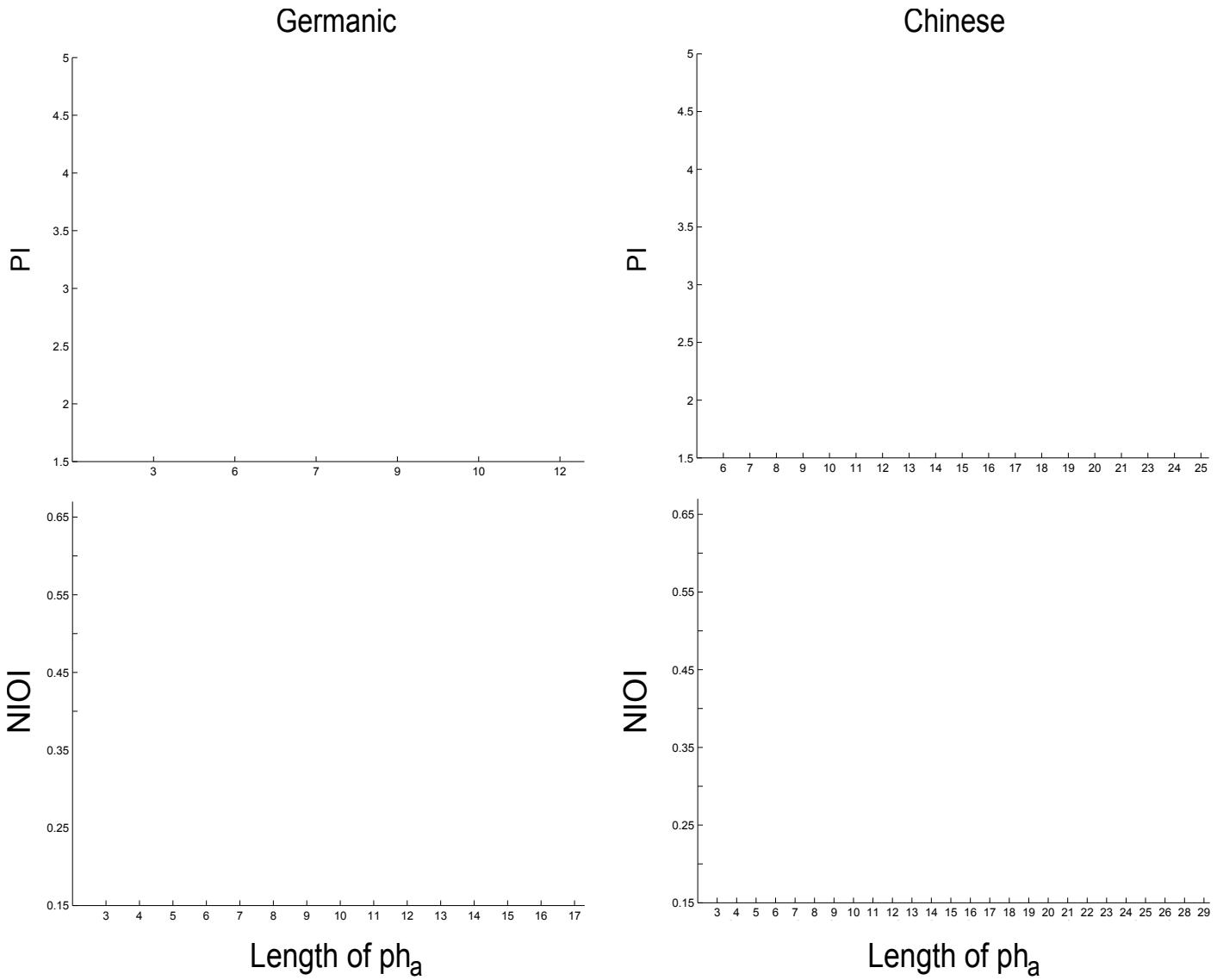

Corpus: Germany | case: All ph_ab | Duration Analysis
N IOI
0.7
0.6
0.5
0.4
0.3
0.2
0.1
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals


Corpus: China | case: All ph_ab | Duration Analysis
N IOI
0.8
0.6
0.4
0.2
J(pab)
C(pa)
C(pb)
Join Interval and Context Intervals

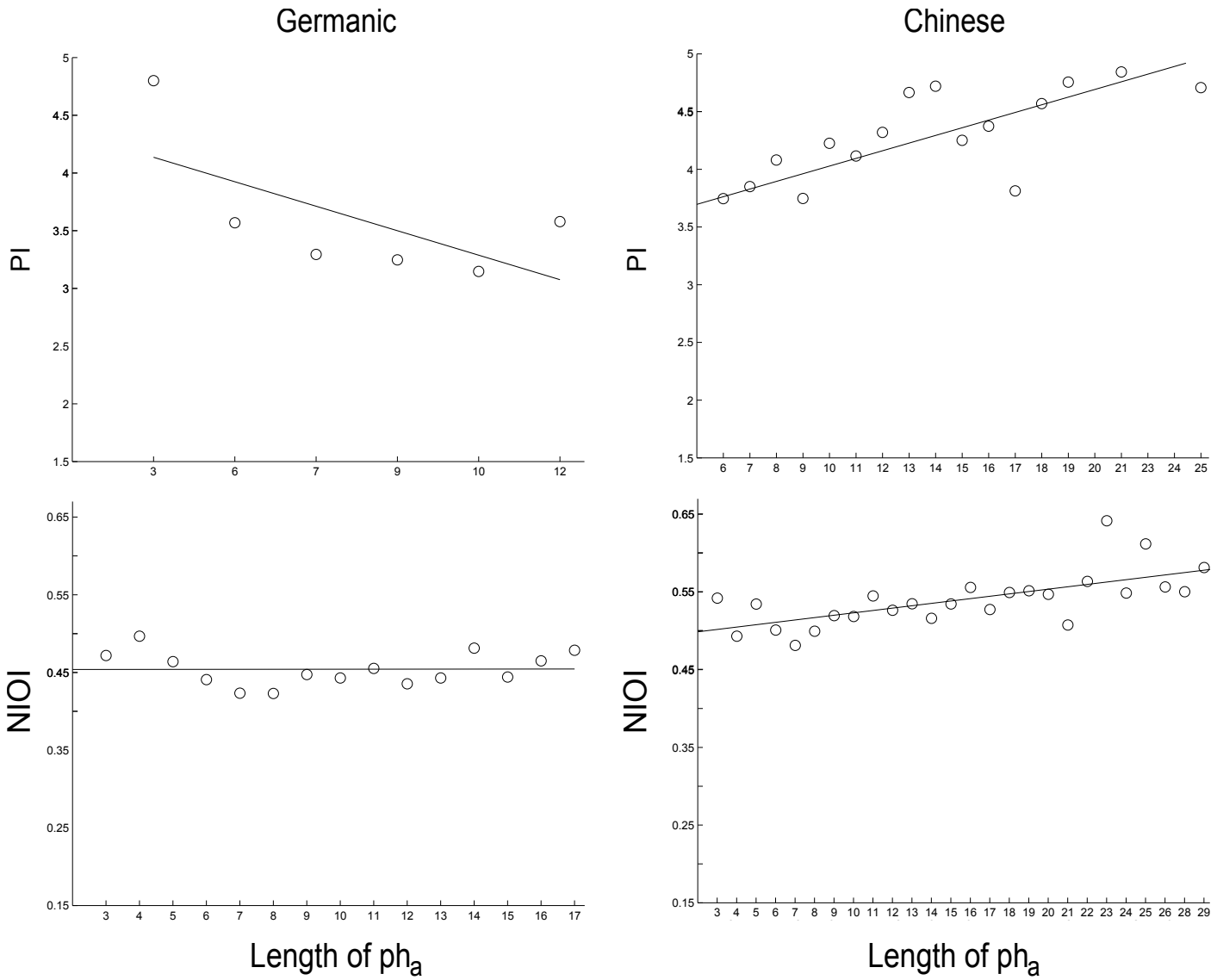
Results 3:

- group phrase pairs in respect to size(ph_a)
- pairwise t-test between m-j($ph_{a,b}$) and m-c($ph_{a,b}$)
- **China** PI: 19/26, NIOI: 26/26
- **Germany** PI: 6/19, NIOI: 15/19

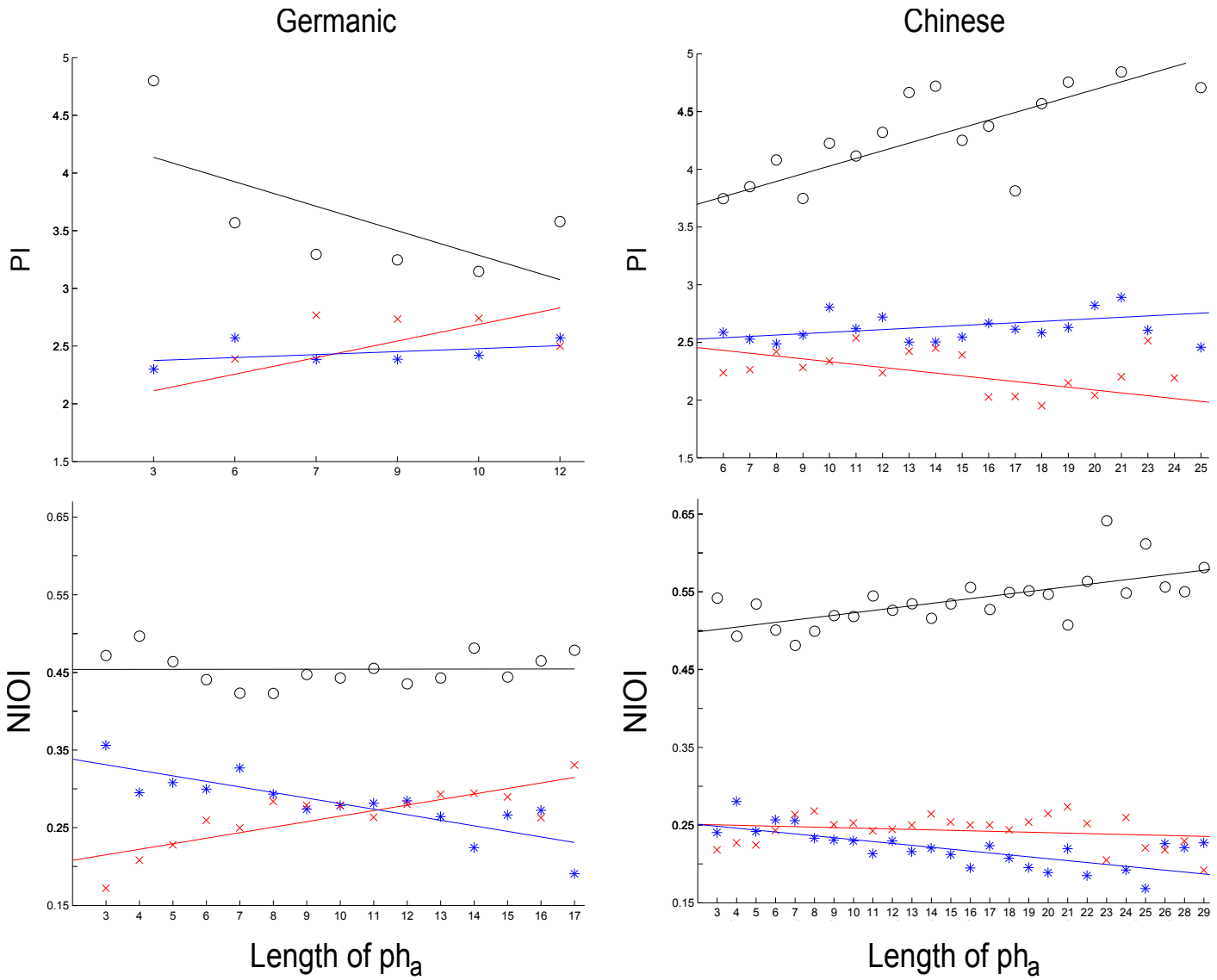
Results 3:



Results 3:



Results 3:



Conclusions

Under the scope of our study:

- discontinuities in **pitch**
  - are **weak** predictors of melodic phrase boundaries
  - **can not** be considered universal
- discontinuities in **duration**
  - are **strong** predictors of phrase boundaries
  - **remain stable** independent of cultural origin and size

Universiteit Utrecht

NWO
Netherlands Organisation for Scientific Research

THANK YOU!