Articulatory discriminability of vowels: Articulator and corpus effects

Philip Hoole

Inst. für Phonetik, Munich Univ., Schellingstr. 3,
D-80799 Munich, Germany,
hoole@phonetik.uni-muenchen.de

137th ASA meeting. Berlin March 14-19, 1999. Session 5pSCb5. : Friday Afternoon, March 19

AIMS

Main aim:

What contribution do different sources of articulatory information make to the discriminability of vowels?

Secondary aim:

How do different types of corpora affect the results of articulatory analysis?

Here we compare real-word vs. highly-constrained pseudo-word corpora from two points of view:

1. In the discriminant analyses used to compare different sources of articulatory information
2. In PARAFAC factor analyses of tongue configuration

ARTICULATORY DATA

Data was acquired with Electromagnetic Articulography (see Fig. 1)

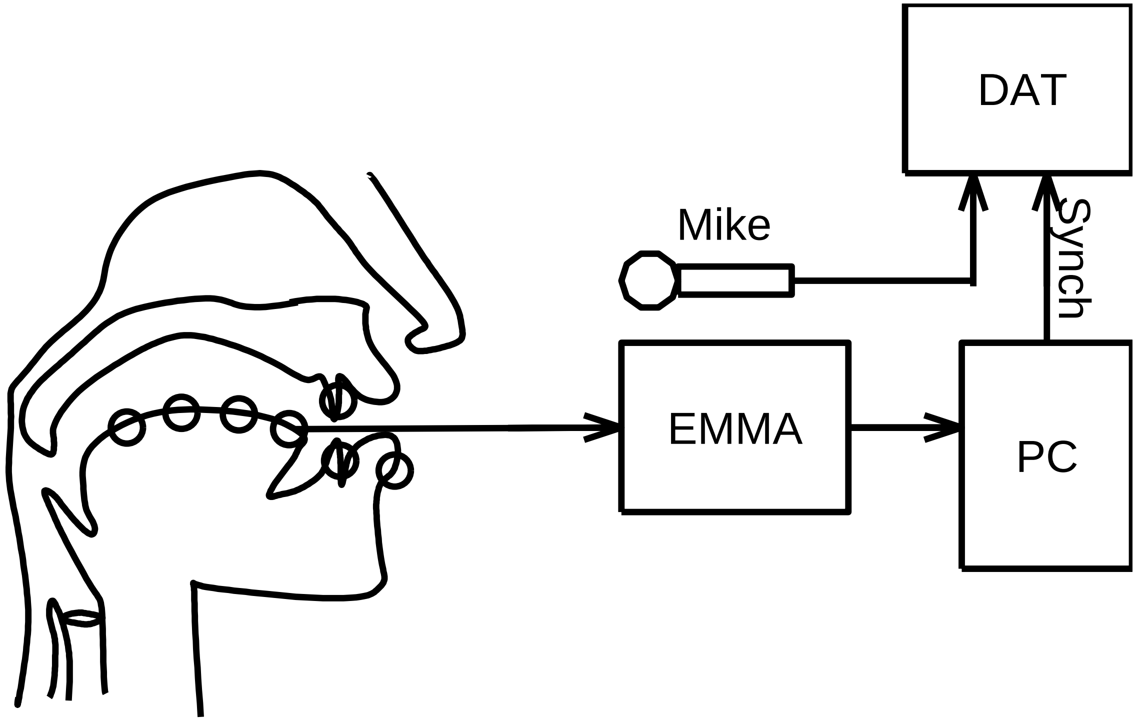


Fig. 1

Tongue	4 sensors
Lower Lip	1 sensor
Jaw	1 sensor

The position, velocity and acceleration of each sensor was extracted at the midpoint of the target vowels.

7 speakers

CORPORA

Corpus A: Pseudo-word, normal speech rate

Corpus B: Pseudo-word, fast speech rate

Corpus C: Real words

All 3 corpora contained the following target vowels:

V = **/ iː ɪ yː ʏ eː ɛ ɛː øː œ ɑː a oː ɔ uː ʊ /**

Corpus A and **Corpus B** embedded the target vowels in a constant carrier phrase

"Ich habe /gəC_1VC_2ə/ gesagt"

with $C_1 = C_2$ = /p, t, k/

5 repetitions of each CV combination

Corpus C embedded each target vowel in 15 different real words, embedded in turn in non-stereotyped sentences

e.g “Der <u>n**e**tte</u> <u>Sch**o**tte</u> hat eine schwarze <u>S**o**cke</u> verloren”

“Eine <u>M**o**tte</u> und eine <u>M**ü**cke</u> klebten am Fliegenpapier”

(target word underlined, target vowel bold face)

DISCRIMINANT ANALYSES

Separate discriminant analyses were run for each of the seven speakers and each corpus.
Only the phonemic identity of the vowel was used for classification, i.e the vowels were not subcategorized with respect to consonant context. Thus each corpus included 15 tokens of each vowel.
The analyses were run for 6 selected combinations of articulatory parameters: The first four used only positional information. The last two additionally used velocity information (acceleration gave similar results). The first condition (i.e position data from the 4 tongue sensors) was used as a baseline condition.

	Articulators	Data category	Abbreviation
1.	Tongue	Position	(T.p)
2.	Tongue and Jaw	Position	(TJ.p)
3.	Tongue and Lip	Position	(TL.p)
4.	Tongue, Jaw and Lip	Position	(TJL.p)
5.	Tongue	Position and Velocity	(T.pv)
6.	Tongue, Jaw and Lip	Position and Velocity	(TJL.pv)

Results are presented throughout as percent correct classification, averaged over speakers.

Results

Overall (Fig. 2)

Parameter combinations are arranged essentially in increasing order of classification accuracy from T.p (tongue; position) on the left, to TJL.pv (Tongue, Jaw and Lip; Position and Velocity) on the right (the exception to this arrangement is that the two Position and Velocity parameter combinations have been kept together).

Relative to the T.p baseline, adding in additional articulatory information results in the following approximate classification increments:

Jaw Position	9%
Lip Position	14%
Jaw and Lip Position	17%
Position **and** Velocity for all articulators	23%.

The three corpora show very similar patterns, simply differing in overall accuracy.
Corpus C (real words) is consistently about 10% worse than Corpus A (pseudo-words, normal rate), with Corpus B (pseudo-word, fast rate) in between.

Details (Figs. 3-5)

Inspection of the results for individual vowels suggested the formation of 6 groups: *front unrounded*, *front rounded*, *back*; each divided into *tense* and *lax*.

- *Back tense* are overall classified best, but increase comparatively little from the T.p baseline condition.
- *Lax* vowels (right panels) are classified worse than tense (left panels), but improve more sharply from the baseline condition.

  There is thus less difference between vowels and corpora in the maximum-information TJL.pv condition.

  Exception: Classification of both *front lax* groups remains noticeably poor for Corpus C in the TJL.pv condition.
- All groups except *back tense* profit strongly from lip information. Contrast e.g. T.p with TL.p, or T.pv with TJL.pv.

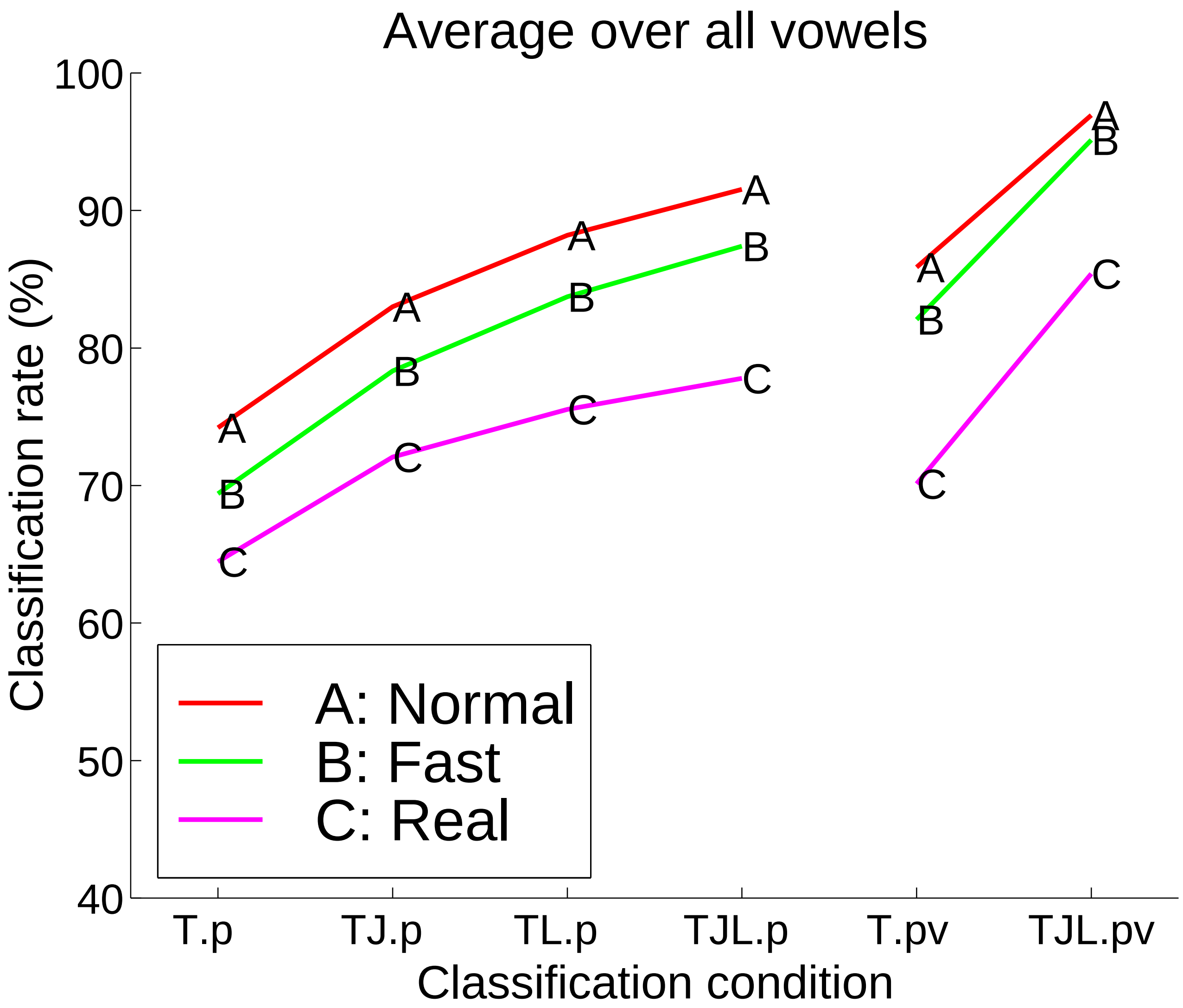
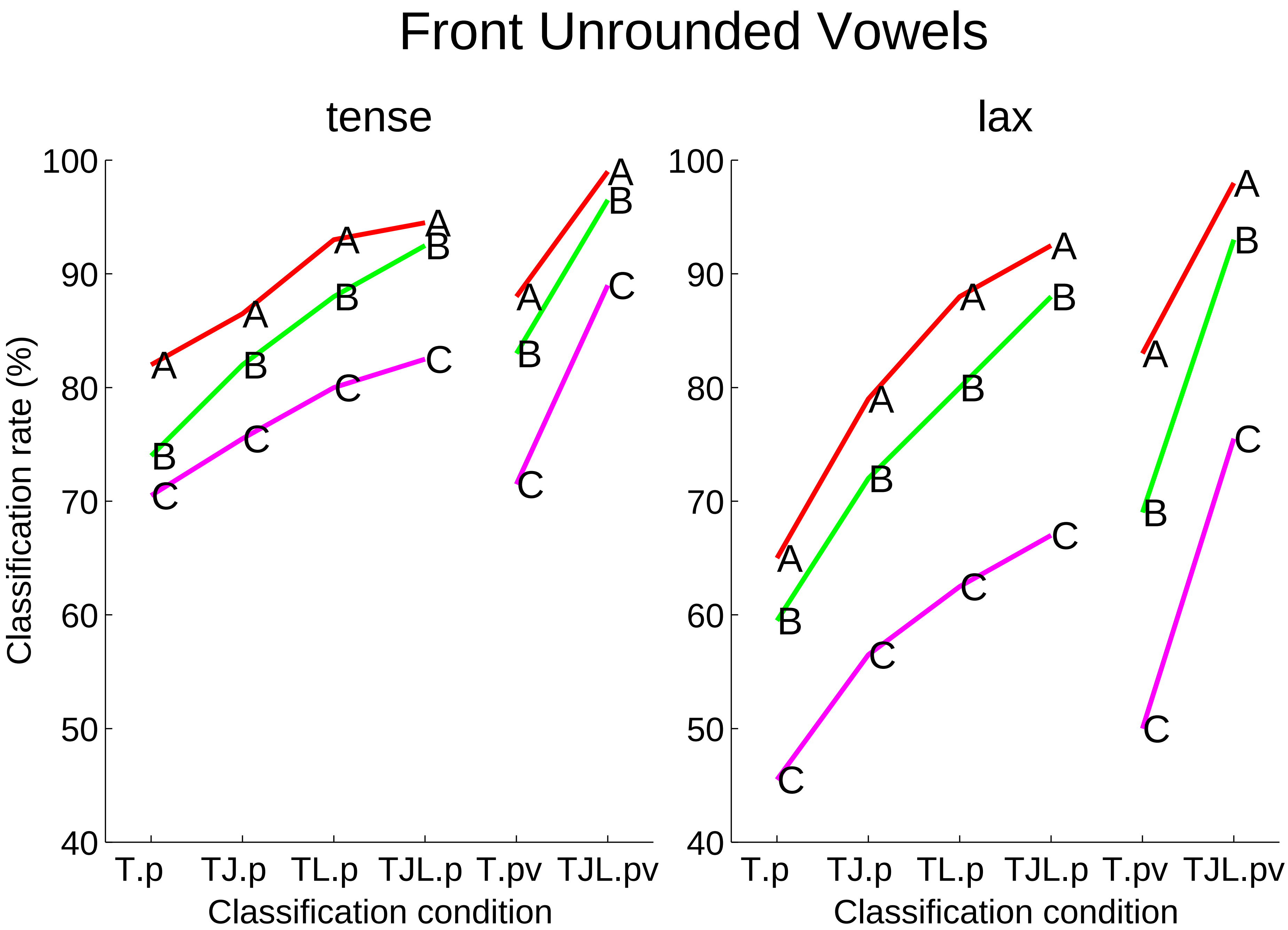


Fig. 2. Dependency of vowel classification on articulatory parameter combination and on corpus. The four leftmost conditions use only position information; the two rightmost conditions use position and velocity information.



Front Unrounded Vowels
tense
lax
Classification rate (%)
100
90
80
70
60
50
40
T.p
TJ.p
TL.p
TJL.p
T.pv
TJL.pv
Classification condition
A
B
C

Fig. 3

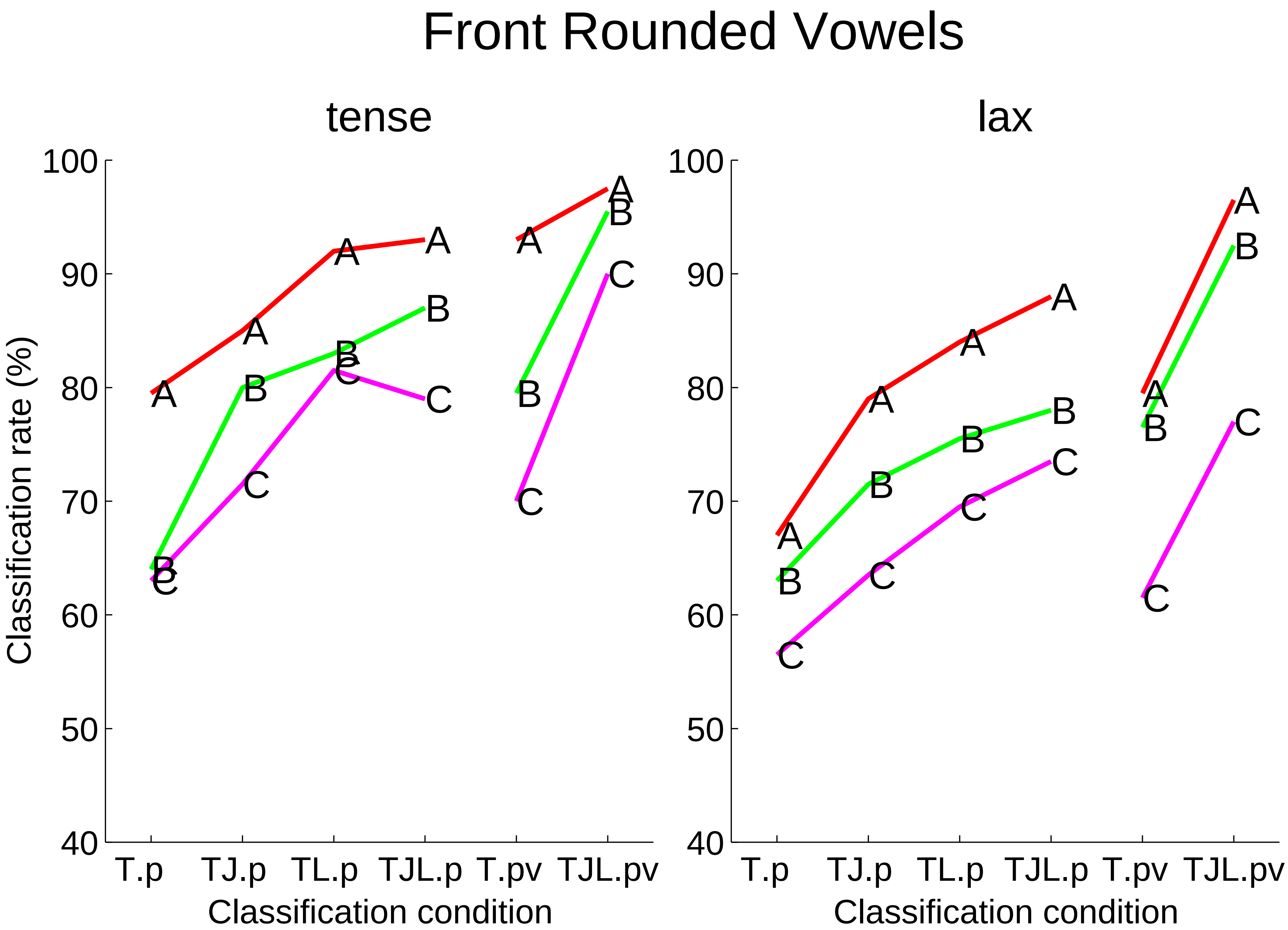
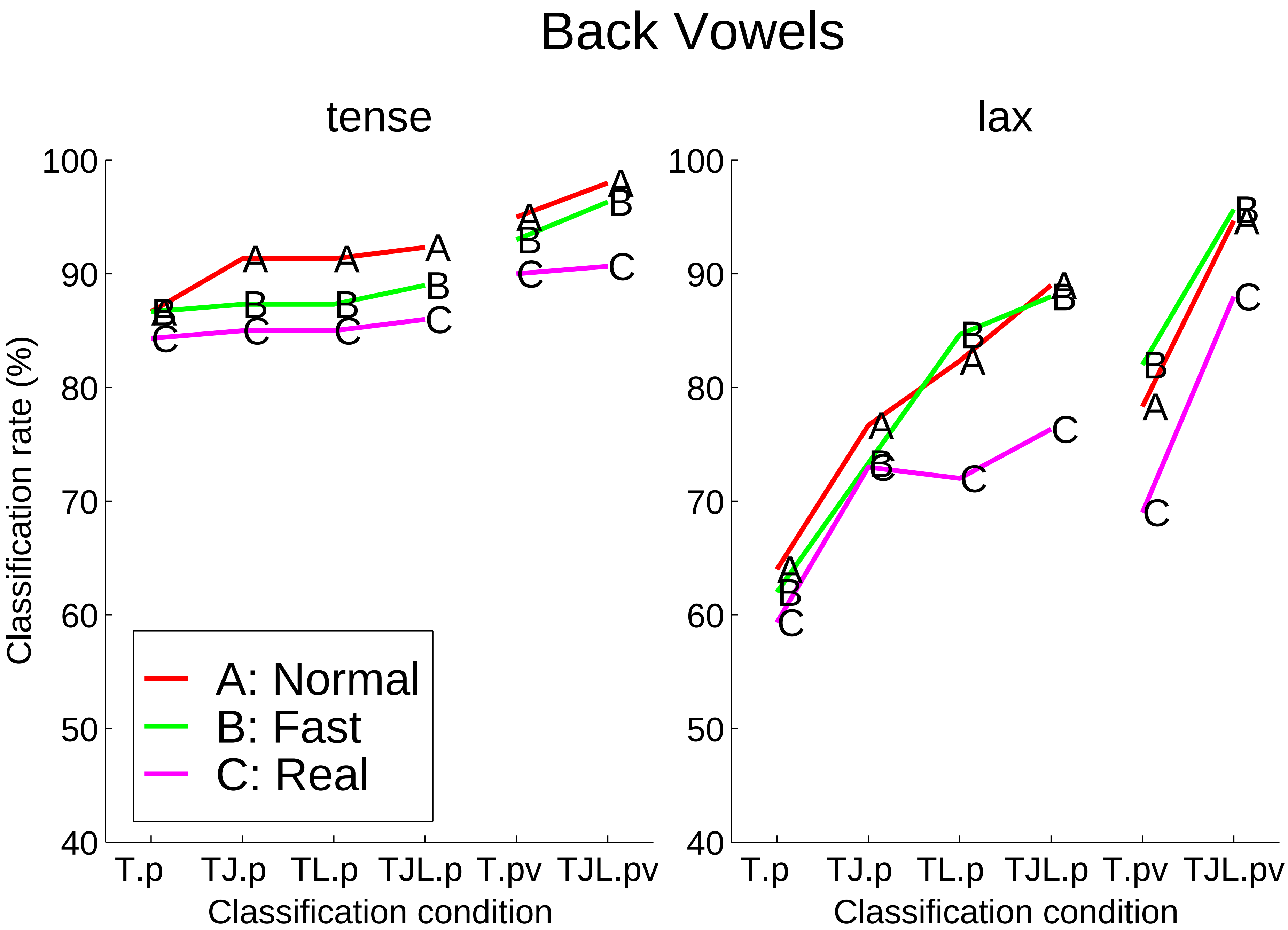

Front Rounded Vowels
tense
lax
Classification rate (%)
100
90
80
70
60
50
40
T.p
TJ.p
TL.p
TJL.p
T.pv
TJL.pv
Classification condition
A
B
C

Fig. 4



Back Vowels
tense
lax
Classification rate (%)
100
90
80
70
60
50
40
T.p
TJ.p
TL.p
TJL.p
T.pv
TJL.pv
Classification condition
A: Normal
B: Fast
C: Real

Fig. 5

PARAFAC Factor Analysis

PARAFAC can be a useful technique for uncovering underlying patterns of articulatory organization in multi-speaker datasets [1].

In [2] we discuss the extraction of a two-factor PARAFAC model of tongue configurations in vowel production for the two pseudo-word Corpora A and B.

Fig. 6 shows the two families of tongue shapes associated with the two factors of the model.

The left panel of Fig. 8 shows the German vowels in the articulatory vowel space defined by these factors.

Fig. 7 and the right panel of Fig. 8 show the result of extracting a new two-factor model based on the real-word Corpus C.

Although there is a slight shift in the orientation of the factor space the solutions appear very similar:

Factor 1 captures a contrast between low back and high front vowels.

Factor 2 captures a contrast between mid front and high back vowels.

The arrangement of the vowels in the articulatory vowel space is also very similar for both corpus types. Compare, for example the relative positions of /e:/, /y:/ and /ɪ/.

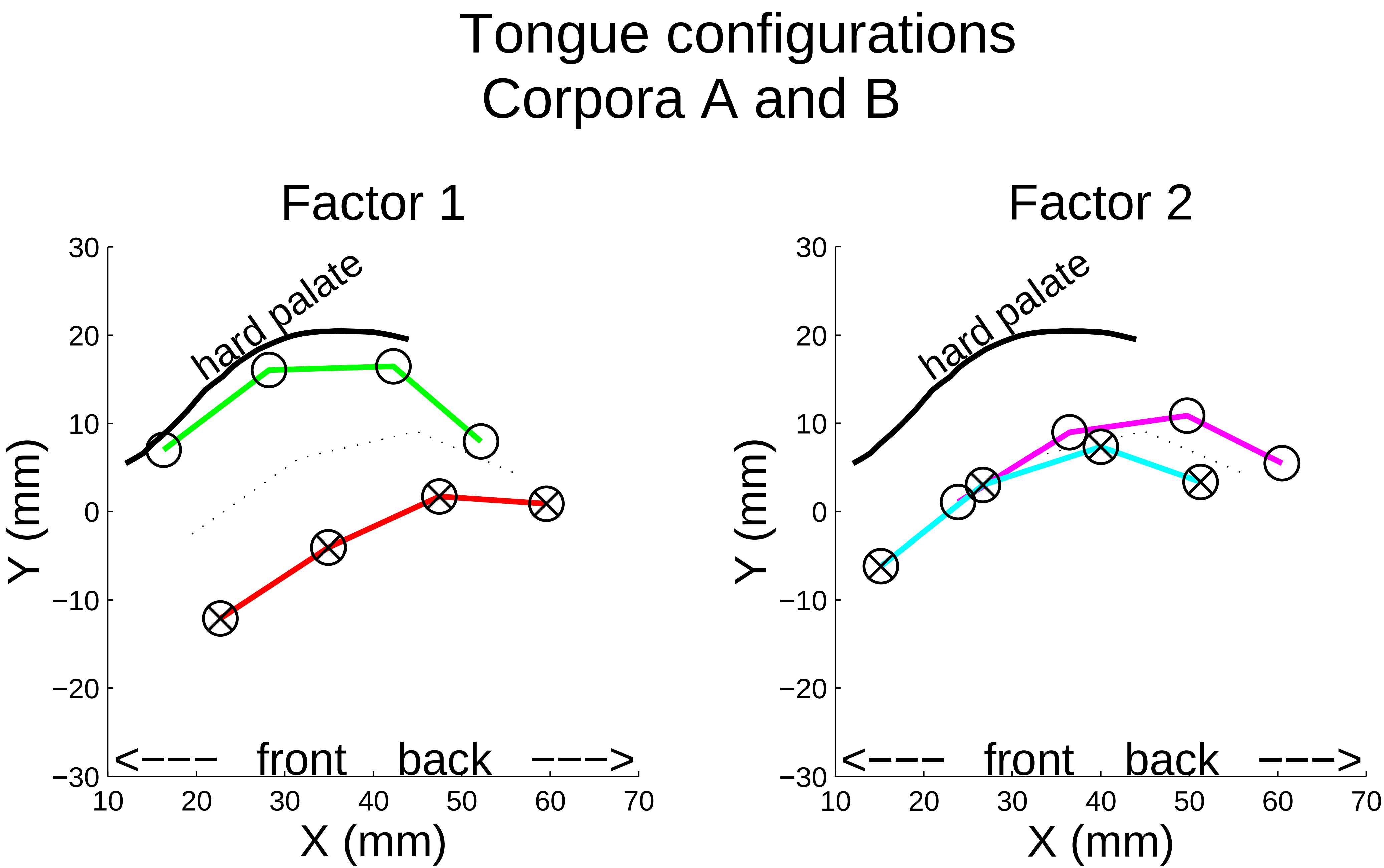
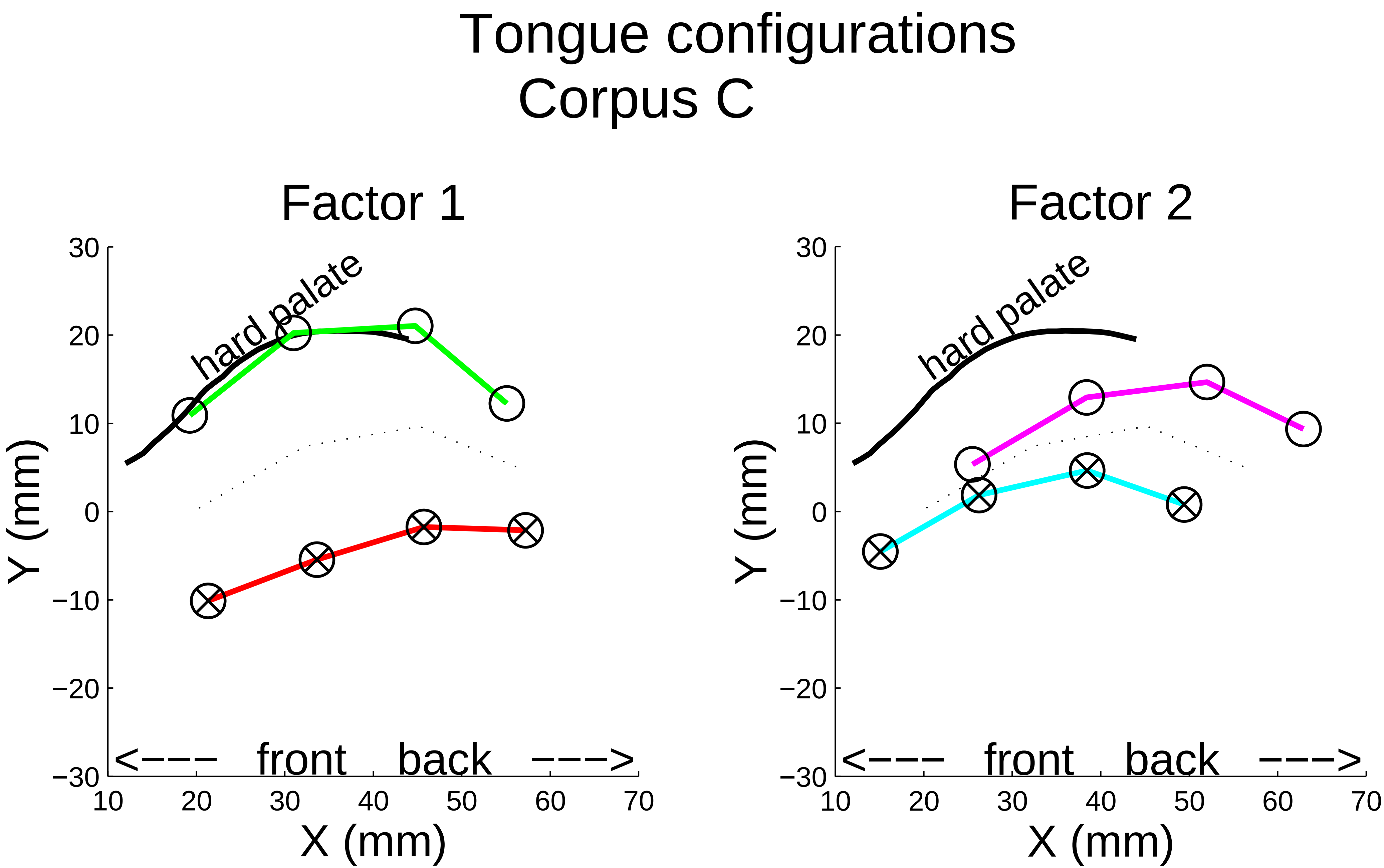

Tongue configurations
Corpora A and B
Factor 1
Factor 2
hard palate
hard palate
Y (mm)
Y (mm)
X (mm)
X (mm)
<---- front back --->
<---- front back --->
30
20
10
0
−10
−20
−30
10 20 30 40 50 60 70

Fig. 6



Tongue configurations
Corpus C
Factor 1
Factor 2
hard palate
hard palate
Y (mm)
X (mm)
<--- front back --->
30
20
10
0
−10
−20
−30
10 20 30 40 50 60 70

Fig. 7

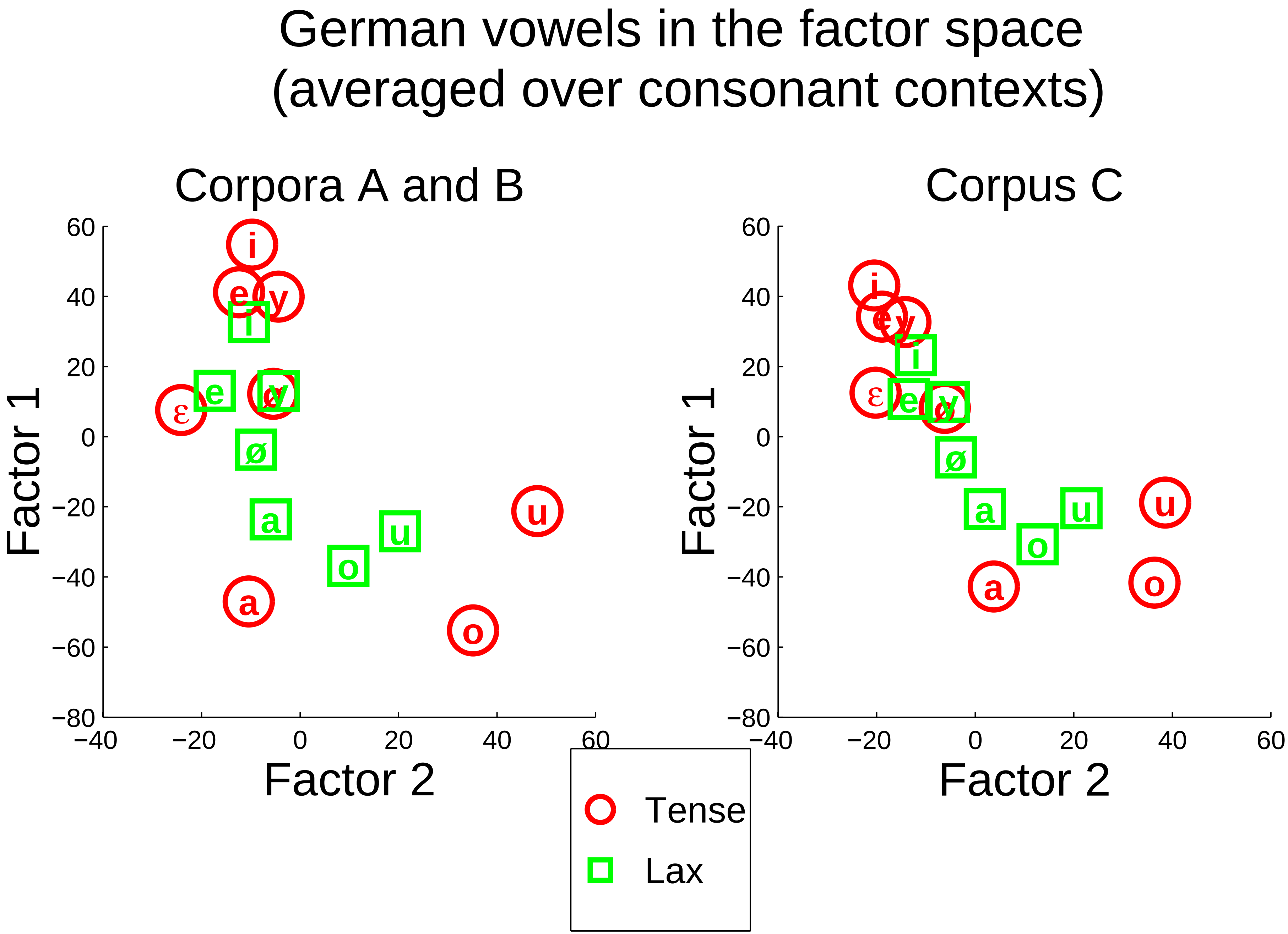

German vowels in the factor space
(averaged over consonant contexts)
Corpora A and B
Corpus C
Factor 1
Factor 2
Tense
Lax
60
40
20
0
−20
−40
−60
−80
−40
−20
0
20
40
60
i
e
y
ε
ø
a
o
u

Fig. 8

SUMMARY and CONCLUSIONS

- Not surprisingly, the classification results can be closely related to the distribution of the vowels in the German vowel space. Back tense vowels have fewest close neighbours, while lax front vowels have most (cf. Fig. 8). Accordingly, these are the two vowel groups whose results contrast most sharply.

  Although back tense vowels are classified very reliably, they actually have high levels of articulatory variability (see discussion in [3]).
- Even though front rounded and unrounded vowel pairs consistently differ in tongue position (Fig. 8), labial information is still crucial to discrimination in the high front vowel space.
- The pseudo-word and the real-word corpora showed very similar patterns, both in the discriminance and in the PARAFAC analyses.

  Exception: The results for the pseudo-word corpora, if taken in isolation, could have suggested that good discrimination can be achieved without taking contextual effects into account. The poor discrimination of front lax vowels in the more realistic Corpus C shows this would be a misleading conclusion.

References

[1] Harshman, R., Ladefoged, P. & Goldstein, L. (1977). "*Factor Analysis of Tongue Shapes*," J. Acoust. Soc. Am. 62, 693- 707.

[2] Hoole, P. "*Modelling tongue position in German vowels*", J. Acoust. Soc. Am. (submitted).

[3] Hoole, P. & Kühnert, B. (1995). "*Patterns of lingual variability in German vowel production*," Proceedings XIIIth Int. Conf. Phon. Sci., Stockholm 2, 442-445.