

On the generalizability of Pāṇini's *pratyāhāra*-technique to other languages

Wiebke Petersen Silke Hamann

Heinrich-Heine University Düsseldorf (Germany)

4th Int. Sanskrit Computational Linguistics Symposium, Dec. 10–12 2010

अइउण्। ऋलृक्। एओङ्। ऐऔच्। हयवरट्।
लण्। ञमङणनम्। झभञ्। घढधष्। जबगडदश्।
खफछठथचटतव्। कपय्। शषसर्। हल्।

Pāṇini's Śivasūtras

1.	a	i	u			N	अइउण्। ऋलृक्।
2.				r	!	K	<i>a·i·uṅ ṛ·lṛk </i>
3.		e	o			Ñ	एओङ्। ऐऔच्।
4.		ai	au			C	<i>e·oñ ai·auc </i>
5.	h	y	v	r		T	हयवरट्। लण्।
6.					l	N	<i>hayavarat laṅ </i>
7.	ñ	m	ṅ	ṇ	n	M	ञमङणनम्। झभञ्।
8.	jh	bh				Ñ	<i>ñamaṅaṅanam jhabhañ </i>
9.			gh	ḍh	dh	Ṣ	घढधष्। जबगडदश्।
10.	j	b	g	ḍ	d	Ṣ	<i>ghadhadhaṣ jabagadadaś </i>
11.	kh	ph	ch	ṭh	th		खफछठथचटतव्।
			c	ṭ	t	V	<i>khaphachathathacaṭatav </i>
12.	k	p				Y	कपय्। शषसर्। हल्।
13.			ś	ṣ	s	R	<i>kapay śaṣasar hal </i>
14.	h					L	

Pāṇini's Śivasūtras

1.	a	i	u			N
2.				r	!	K
3.		e	o			Ñ
4.		ai	au			C
5.	h	y	v	r		T
6.					l	N
7.	ñ	m	ṅ	ṇ	n	M
8.	jh	bh				Ñ̃
9.			gh	ḍh	dh	Ṣ
10.	j	b	g	ḍ	d	Ṣ̃
11.	kh	ph	ch	ṭh	th	
			c	ṭ	t	V
12.	k	p				Y
13.			ś	ṣ	s	R
14.	h					L

अइउण्। ऋलृक्।

a·i·uṅ | ṛ·lṛk

एओङ्। ऐऔच्।

e·oṅ | ai·auc

हयवरट्। लण्।

hayavarat | laṅ

जमङणनम्। झभञ्।

ñamaṅṅanam | jhabhañ

घढधष्। जबगडदश्।

ghadhadhaṣ | jabagaḍadaś

खफछठथचटतव्।

khaphachathathacaṭatav

कपय्। शषसर्। हल्।

kapay | śaṣasar | hal

Pāṇini's Śivasūtras

1.	a	i	u			N
2.				r	!	K
3.		e	o			Ñ
4.		ai	au			C
5.	h	y	v	r		T
6.					l	N
7.	ñ	m	ṅ	ṇ	n	M
8.	jh	bh				Ñ
9.			gh	ḍh	dh	Ṣ
10.	j	b	g	ḍ	d	Ṣ
11.	kh	ph	ch	ṭh	th	
			c	ṭ	t	V
12.	k	p				Y
13.			ś	ṣ	s	R
14.	h					L

अइउण्। ऋलृक्।

a·i·uṅ | ṛ·lṛk

एओङ्। ऐऔच्।

e·oṅ | ai·auc

हयवरट्। लण्।

hayavarat | laṅ

जमङणनम्। झभञ्।

ñamaṅṅanam | jhabhañ

घढधष्। जबगडदश्।

ghadhadhaṣ | jabagaḍadaś

खफछठथचटतव्।

khaphachathathacaṭatav

कपय्। शषसर्। हल्।

kapay | śaṣasar | hal

Pratyāhāras

1.		a	i	u			Ṇ
2.					r	!	Ḳ
3.			e	o			Ṇ̇
4.			ai	au			C
5.		h	y	v	r		Ṭ

Pratyāhāras

1.		a	i	u				Ṇ
2.					r	!		Ḳ
3.			e	o				Ṇ̇
4.			ai	au				Ḳ
5.		h	y	v	r			Ṭ
								aḲ

Pratyāhāras

1.	a	i	u		Ṇ
2.				ṛ	Ḍ
3.		e	o		ṅ
4.		ai	au		Ḍ
5.	h	y	v	r	Ṛ

aC= ⟨a, i, u, ṛ, Ḍ, e, o, ai, au⟩

Analysis of iko yaṇaci: [iK] → [yN]/_ [aC]

1.	a	i	u			Ṅ
2.				ṛ	!	K
3.		e	o			Ṅ
4.		ai	au			C
5.	h	y	v	r		Ṭ
6.					l	Ṅ

- *sūtra* 6.1.77: *iko yaṇaci* (इको यणचि)
- analysis: [ik]_{gen}[yaṇ]_{nom}[ac]_{loc}
- [iK] → [yN]/_ [aC]
- ⟨i, u, ṛ, !⟩ → ⟨y, v, r, l⟩/_ ⟨a, i, u, ṛ, !, e, o, ai, au⟩
- dadhi + atra → dadhyatra

Our task

Pāṇini's method

- Unique method which is especially economical
- Method developed for a single language
- Testing of applicability to other languages

German as a test case

- Indoeuropean language with a smaller Sandhi-system
- No codified grammatical description
- 7 example processes of German

Phonological processes of German

final devoicing /hʊnd/ 'dog' → [hʊnt]

⟨b, d, g, v, z, ʒ⟩ → ⟨p, t, k, f, s, ʃ⟩ / __ word boundary

umlaut [bal+ə] 'ball' (sg.) + plural → [bɛlə]

⟨u, ū, o, ɔ, a, a, aṽ⟩ → ⟨y, ʏ, ø, œ, ε, ε, ɔ̃⟩

velar nasal assimilation /ankʊnft/ 'arrival' → [aŋkʊnft]

⟨n⟩ → ⟨ŋ⟩ / __ ⟨k, g⟩

labial nasal assimilation /anbaʊ/ 'outbuilding' → [ambaʊ]

⟨n⟩ → ⟨m⟩ / __ ⟨p, b, p̂f, f, v⟩

glottal stop epenthesis /aɪ/ 'egg' → [ʔaɪ]

⟨⟩ → ⟨ʔ⟩ / word boundary __ vowel

vowel shortening /mo:næçɪ:/ 'monarchy' → [monæçɪ:]

⟨i, y, e, ø, u, o⟩ → ⟨i, y, e, ø, u, o⟩ / __ unstressed

palatal fricative assimilation /baç/ 'brook' → [bax]

⟨ç⟩ → ⟨x⟩ / ⟨a, ɔ, a, u, ū, o, ɔ, aṽ⟩ __

Set of 10 German sound classes that have to be described

devoicing input {b, d, g, v, z, ʒ}

devoicing output {p, t, k, f, s, ʃ}

umlaut input {uɪ, ʊ, oɪ, ɔ, aɪ, a, aʊ}

umlaut output {yɪ, ʏ, øɪ, œ, εɪ, ε, ɔɪ}

nasal assimilation velar (right context) {k, g}

nasal assimilation labial (right context) {p, b, p̂f, f, v}

glottal stop epenthesis (right context)

{ʏ, ø, ə, ɪ, ε, œ, ɔ, ʊ, a, i, e, o, u, y, ø, aɪ, iɪ, oɪ, uɪ, yɪ, øɪ, eɪ, aɪ, ɔɪ, aʊ, εɪ}

vowel shortening input {iɪ, yɪ, eɪ, øɪ, uɪ, oɪ}

vowel shortening output {i, y, e, ø, u, o}

palatal fricative assimilation {aɪ, a, uɪ, ʊ, oɪ, ɔ, aʊ}

Main theorem on *pratyāhārasūtras*

task

Given a set of sound classes find a list of the sounds in the style of Pāṇini's *Śivasūtras* (called *pratyāhārasūtras*) which provides a *pratyāhāra* for each sound class.

main theorem

A set of sound classes has *pratyāhārasūtras* in which no sound occurs more than once

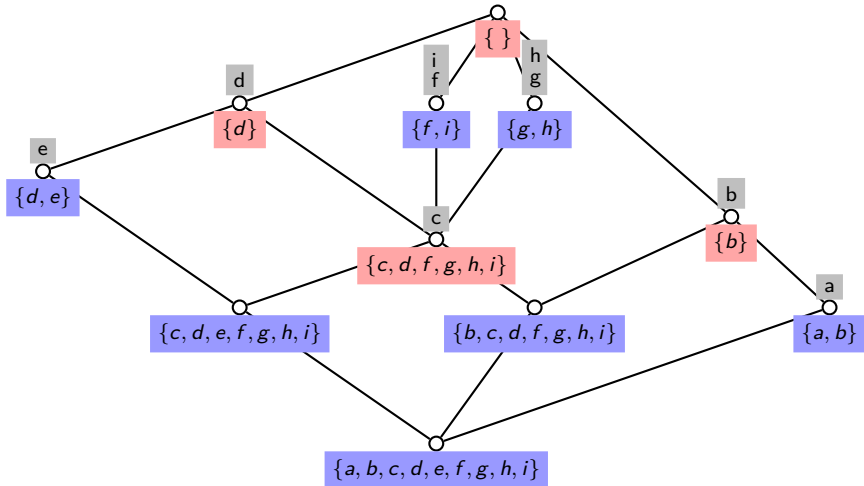
⇔

the **concept lattice** of the set of sound classes is planar and each sound occurs on its *boundary*.

Petersen (2009): On the Construction of *Śivasūtra*-Alphabets. In Proc. of the 3rd Int. Sanskrit Comp. Ling. Symp.

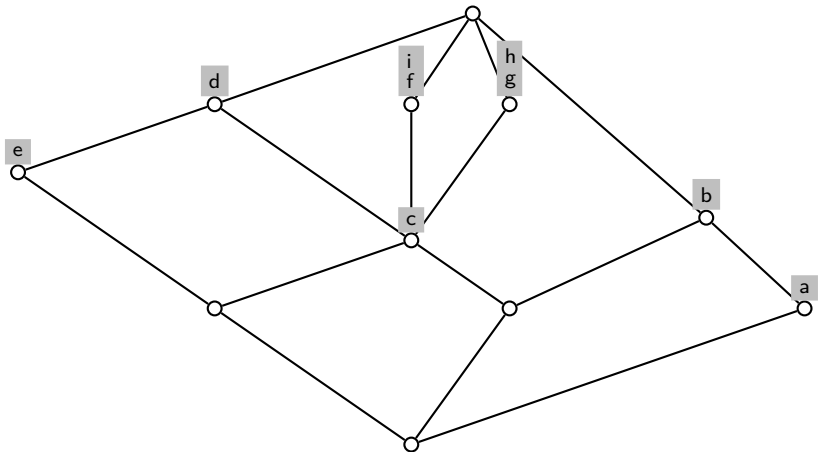
Example: concept lattice

$\{\{d, e\}, \{a, b\}, \{b, c, d, f, g, h, i\}, \{f, i\}, \{c, d, e, f, g, h, i\}, \{g, h\}\}$

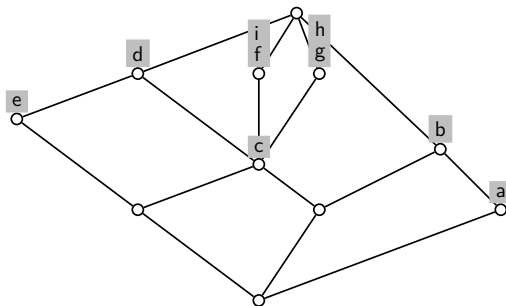


Example: concept lattice

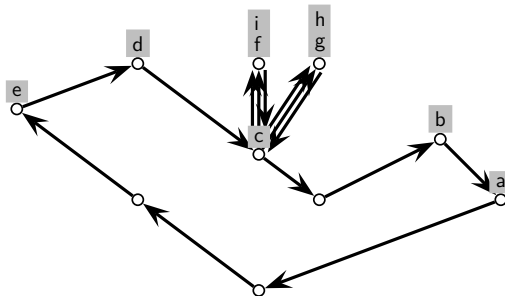
$\{\{d, e\}, \{a, b\}, \{b, c, d, f, g, h, i\}, \{f, i\}, \{c, d, e, f, g, h, i\}, \{g, h\}\}$



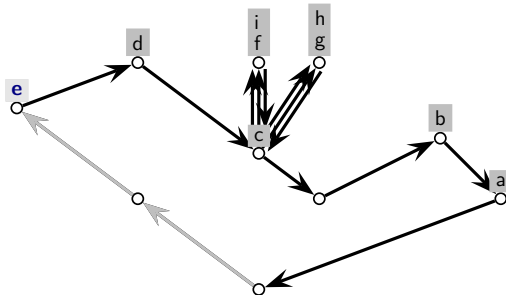
Pratyāhārasūtras with a minimal number of markers



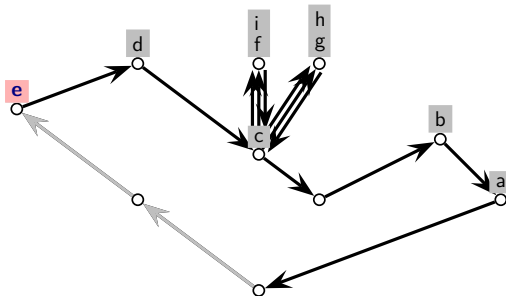
Pratyāhārasūtras with a minimal number of markers



Pratyāhārasūtras with a minimal number of markers

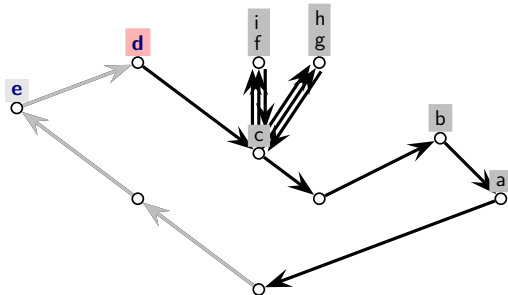


Pratyāhārasūtras with a minimal number of markers



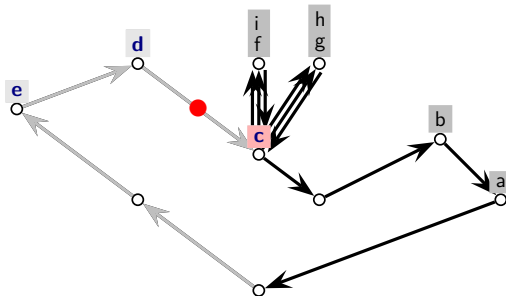
e

Pratyāhārasūtras with a minimal number of markers



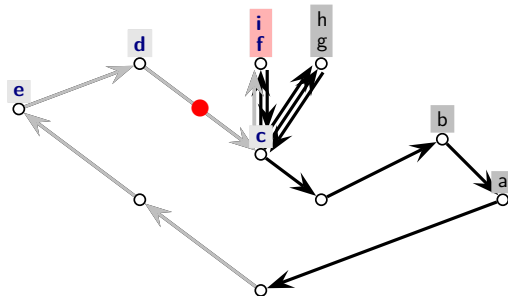
ed

Pratyāhārasūtras with a minimal number of markers



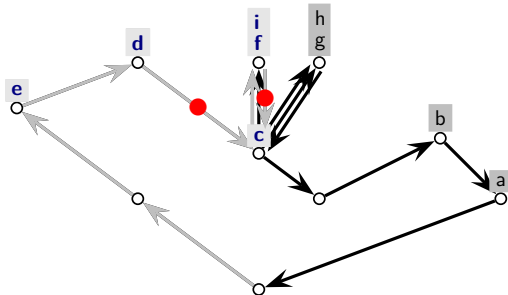
edM_1c

Pratyāhārasūtras with a minimal number of markers



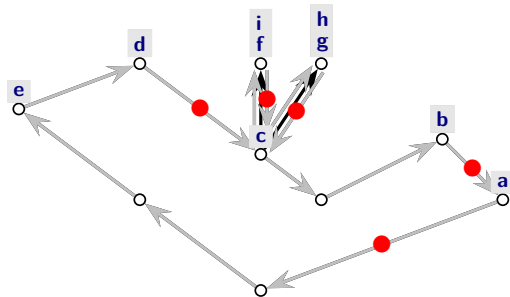
edM_1cfi

Pratyāhārasūtras with a minimal number of markers



edM_1cfiM_2

Pratyāhārasūtras with a minimal number of markers



$edM_1cfiM_2ghM_3bM_4aM_5$

Set of 10 German sound classes that have to be described

devoicing input {b, d, g, v, z, ʒ}

devoicing output {p, t, k, f, s, ʃ}

umlaut input {uɪ, υ, oɪ, ɔ, aɪ, a, aʊ}

umlaut output {yɪ, ʏ, øɪ, œ, εɪ, ε, ɔɪ}

nasal assimilation velar (right context) {k, g}

nasal assimilation labial (right context) {p, b, p̂f, f, v}

glottal stop epenthesis (right context)

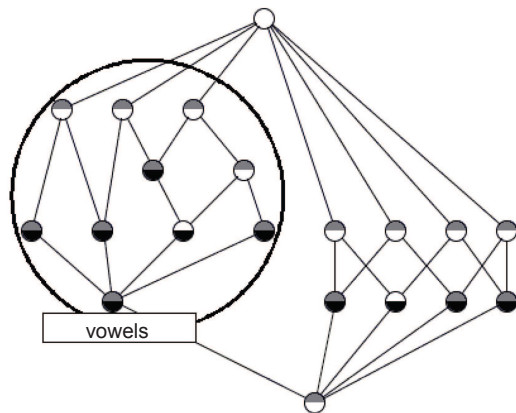
{ʏ, ø, ə, ɪ, ε, œ, ɔ, υ, a, i, e, o, u, y, ø, aɪ, iɪ, oɪ, uɪ, yɪ, øɪ, eɪ, aɪ, ɔɪ, aʊ, εɪ}

vowel shortening input {iɪ, yɪ, eɪ, øɪ, uɪ, oɪ}

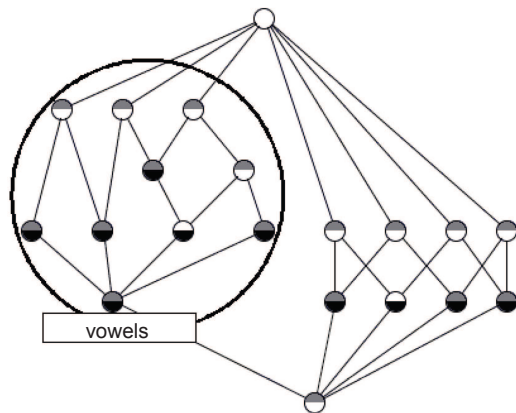
vowel shortening output {i, y, e, ø, u, o}

palatal fricative assimilation {aɪ, a, uɪ, υ, oɪ, ɔ, aʊ}

German sound classes: concept lattice

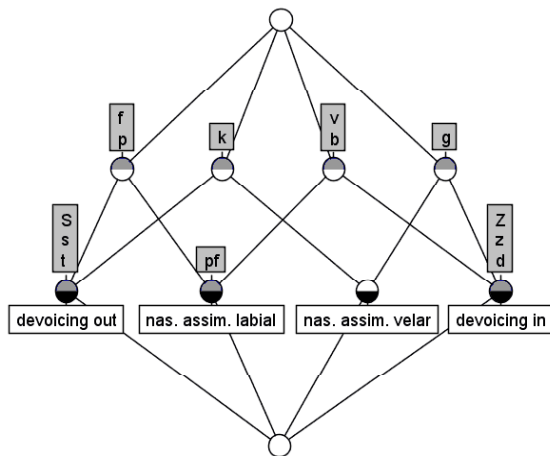


German sound classes: concept lattice

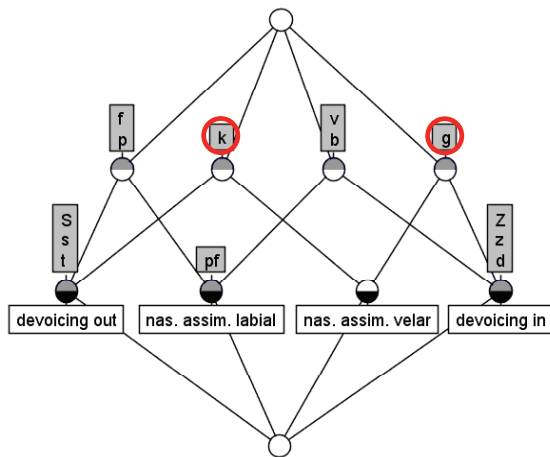


Vowels constitute a planar lattice

Consonants constitute a non-planar lattice

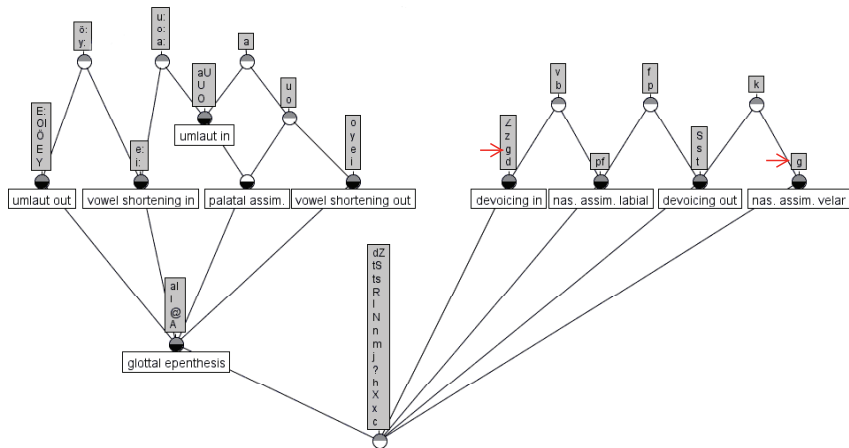


Consonants constitute a non-planar lattice



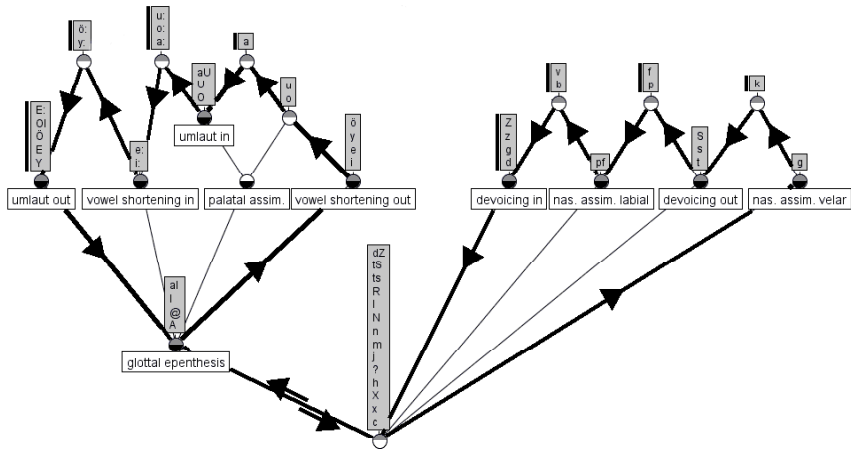
Duplication of a consonant necessary!

Proposed *pratyāhārasūtras* for German



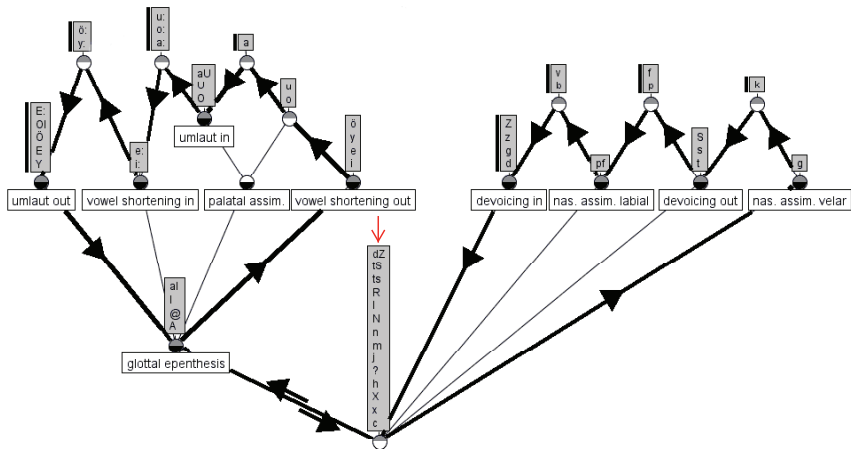
$\widehat{d}\widehat{z} \widehat{t}\widehat{f} \widehat{t}\widehat{s} \text{ ʋ } l \eta n m j ? h \zeta x \text{ g } k M_1 t s f f p M_2 \widehat{p}\widehat{f} b v M_3 \text{ ʒ } z d \text{ g } M_4$
 $\text{ ə } \text{ ɐ } i \text{ a}_i \text{ } \emptyset y e i u o a M_5 \text{ ʊ } \text{ ɔ } \text{ a}_U \text{ a: } o: u: M_6 i: e: y: \text{ } \emptyset: M_7 \text{ ɛ: } \text{ ɔ}_I \text{ } \text{œ } \text{ ʏ } \text{ ɛ } M_8$

Proposed *pratyāhārasūtras* for German



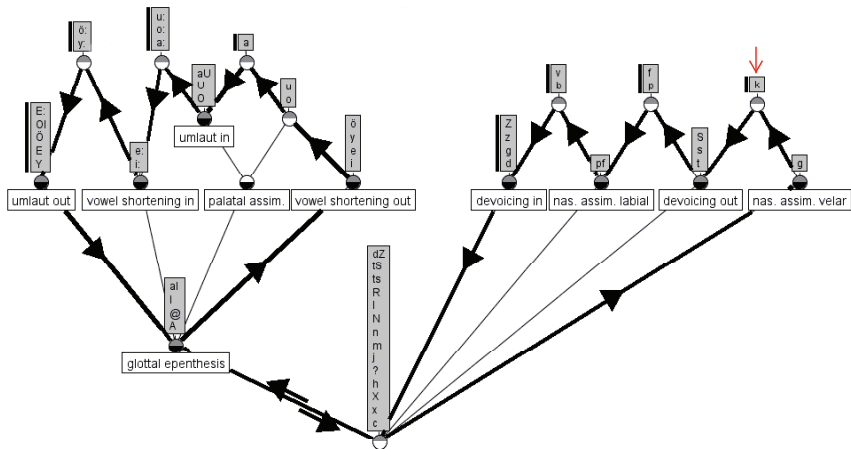
$\widehat{d}_3 \widehat{t}_f \widehat{t}_s \text{ ʃ } l \eta n m j \text{ ʔ } h \text{ ç } x g k M_1 t s \text{ ʃ } f p M_2 \widehat{p}f b v M_3 \text{ ʒ } z d g M_4$
 $\text{ ə } \text{ ɐ } i \text{ a}_i \text{ ø } y e i u o a M_5 \text{ ʊ } \text{ ɔ } \text{ a}_y \text{ a: } o: u: M_6 i: e: y: \text{ ø: } M_7 \text{ ɛ: } \text{ ɔ}_i \text{ œ } \text{ ɤ } \text{ ɛ } M_8$

Proposed *pratyāhārasūtras* for German



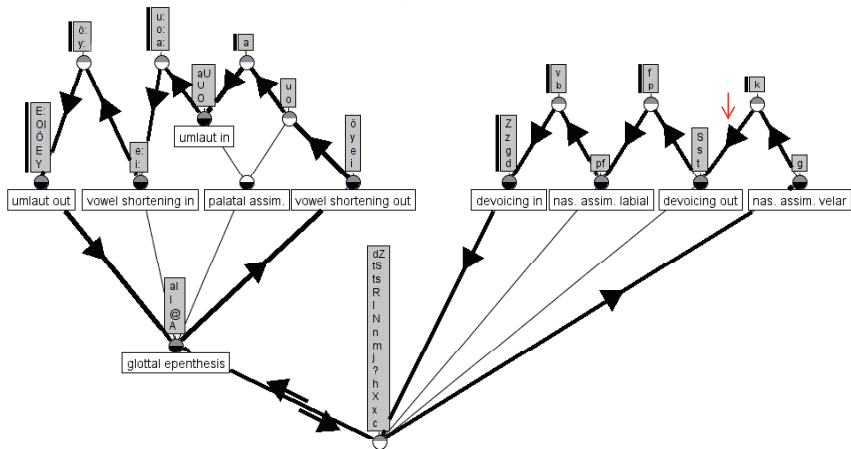
$\widehat{d}_3 \widehat{t}_f \widehat{ts} \text{ ʃ } l \text{ ʎ } n \text{ m } j \text{ ʔ } h \text{ ç } x$ g k M_1 t s ʃ f p M_2 $\widehat{p}f$ b v M_3 ʒ z d g M_4
 ə ɐ i ḁ ø y e i u o a M_5 ʊ ɔ ḁ a: o: u: M_6 i: e: y: ø: M_7 ɛ: ɔ̥ œ ɤ ε M_8

Proposed *pratyāhārasūtras* for German



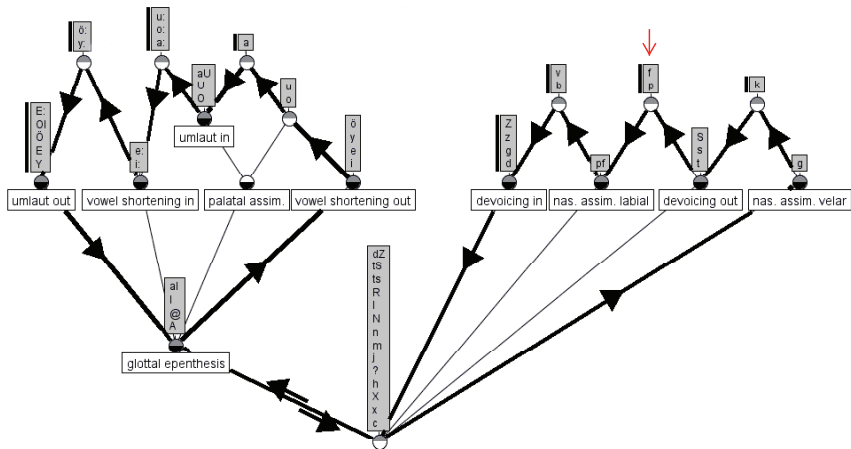
$\widehat{d}_3 \widehat{t}_f \widehat{ts} \text{ ɸ } l \text{ ɲ } n \text{ m } j \text{ ? } h \text{ ç } x \text{ g } k$ M_1 t s f f p M_2 $\widehat{p} \widehat{f}$ b v M_3 ʒ z d g M_4
 ə ɐ i ḁ ø y e i u o a M_5 ʊ ə a̯ a: o: u: M_6 i: e: y: ø: M_7 ɛ: ɔ̥ œ ɤ ε M_8

Proposed *pratyāhārasūtras* for German



$\widehat{d}_3 \widehat{t}_f \widehat{ts} \text{ ɸ } l \text{ ɲ } n \text{ m } j \text{ ? } h \text{ ç } x \text{ g } k \text{ } M_1 \text{ } t \text{ s } f \text{ p } M_2 \text{ } \widehat{p}f \text{ b } v \text{ } M_3 \text{ } ʒ \text{ z } d \text{ g } M_4$
 $\text{ ə } \text{ ɐ } \text{ i } \text{ a}_i \text{ } \emptyset \text{ } y \text{ e } \text{ i } \text{ u } \text{ o } \text{ a } M_5 \text{ } \text{ ʊ } \text{ ɔ } \text{ a}_y \text{ } a \text{ : } o \text{ : } u \text{ : } M_6 \text{ } i \text{ : } e \text{ : } y \text{ : } \emptyset \text{ : } M_7 \text{ } \text{ ɛ } \text{ ɔ}_i \text{ } \text{ œ } \text{ } \text{ ɤ } \text{ } \text{ ɛ } M_8$

Proposed *pratyāhārasūtras* for German



$\widehat{d}_3 \widehat{t}_f \widehat{ts} \text{ ʁ } l \text{ ʃ } n \text{ m } j \text{ ? } h \text{ ç } x \text{ g } k \text{ } M_1 \text{ } t \text{ s } f \text{ f } p \text{ } M_2 \text{ } p \text{ f } b \text{ v } M_3 \text{ } ʒ \text{ z } d \text{ g } M_4$
 $\text{ ə } \text{ ɐ } \text{ ɪ } \text{ a}_i \text{ } \emptyset \text{ } y \text{ e } \text{ i } \text{ u } \text{ o } \text{ a } M_5 \text{ } \text{ ʊ } \text{ ɔ } \text{ a}_y \text{ } a \text{ : } o \text{ : } u \text{ : } M_6 \text{ } i \text{ : } e \text{ : } y \text{ : } \emptyset \text{ : } M_7 \text{ } \text{ ɛ } \text{ ɔ}_i \text{ } \text{ œ } \text{ ɤ } \text{ ɛ } M_8$

Resulting *pratyāhāras*

\widehat{d}_3 \widehat{t}_f \widehat{ts} ɸ l η n m j $?$ h ζ x g k M_1 t s \int f p M_2 $\widehat{p}f$ b v M_3
 ζ z d g M_4 ə ɐ i $a_{\text{̇}}$ ø y e i u o a M_5 ʊ ɔ $a_{\text{̇}}$ $a:$ $o:$ $u:$ M_6 $i:$ $e:$ $y:$ $\text{ø}:$ M_7
 $\varepsilon:$ $\text{ɔ}_{\text{̇}}$ œ ɤ ε M_8

- bM_4 Input to final devoicing
- kM_2 Output to final devoicing
- gM_1 Input to velar nasal assimilation
- fM_3 Input to labial nasal assimilation
- uM_6 Context of palatal fricative assimilation
- aM_6 Input to umlaut
- $y:M_8$ Output to umlaut
- $a:M_7$ Input to vowel shortening
- $\text{ø}M_5$ Output to vowel shortening
- $\text{ə}M_8$ Context of glottal epenthesis

Resulting *pratyāhāras*

\widehat{d}_3 \widehat{t}_f \widehat{t}_s ɸ l η n m j $ʔ$ h ζ x g k M_1 t s \int f p M_2 \widehat{p}_f b v M_3
 ζ z d g M_4 ə ɐ i $a_{\text{̇}}$ ø y e i u o a M_5 ʊ ɔ $a_{\text{̇}}$ $a:$ $o:$ $u:$ M_6 $i:$ $e:$ $y:$ $\text{ø}:$ M_7
 $\text{ɛ}:$ $\text{ɔ}_{\text{̇}}$ œ ɤ ɛ M_8

- bM_4 Input to final devoicing $\langle b v \zeta z d g \rangle$
- kM_2 Output to final devoicing $\langle k t s \int f p \rangle$
- gM_1 Input to velar nasal assimilation
- fM_3 Input to labial nasal assimilation
- uM_6 Context of palatal fricative assimilation
- aM_6 Input to umlaut
- $y:M_8$ Output to umlaut
- $a:M_7$ Input to vowel shortening
- $\text{ø}M_5$ Output to vowel shortening
- $\text{ə}M_8$ Context of glottal epenthesis

Phonological processes in *pratyāhāra* notation

$\widehat{d}_3 \widehat{t}_f \widehat{ts} \text{ ɸ l ŋ n m j } \text{? h } \zeta \text{ x g } \mathbf{k} M_1 \mathbf{t s f f p} M_2 \widehat{p} \mathbf{b v} M_3$
 $\mathbf{ʒ z d g} M_4 \text{ ə ɐ ɪ a}_\text{̣} \text{ ø y e i u o a } M_5 \text{ ʊ ɔ a}_\text{̣} \text{ a: o: u: } M_6 \text{ i: e: y: ø: } M_7$
 $\text{ɛ: ɔ}_\text{̣} \text{ œ } \text{ɣ } \text{ɛ } M_8$

- $bM_4 \rightarrow kM_2 / _ \text{ word boundary}$ *final devoicing*
- $/n/ \rightarrow [ŋ] / _ gM_1$ *velar nasal assimilation*
- $/n/ \rightarrow [m] / _ fM_3$ *labial nasal assimilation*
- $\{ \} \rightarrow [ʔ] / \text{ word boundary } _ \text{ ə}M_8$ *glottal stop epenthesis*
- $a:M_7 \rightarrow \text{ø}M_5 / _ \text{ unstressed}$ *vowel shortening*
- $/\zeta/ \rightarrow [x] / uM_6 _$ *palatal fricative assimilation*
- $aM_6 \rightarrow y:M_8$ *umlaut*

Summary & outlook

- The *pratyāhāra* method is applicable to the description of the sound system of German
- Our sample of German phonological processes is rather small
- Further languages (e.g. non-Indoeuropean, or with larger Sandhi-systems) have to be investigated