

## Phonological overgeneration in Paninian system

Malhar Kulkarni

► **To cite this version:**

Malhar Kulkarni. Phonological overgeneration in Paninian system. Gérard Huet and Amba Kulkarni. First International Sanskrit Computational Linguistics Symposium, Oct 2007, Rocquencourt, France. 2007, <http://hal.inria.fr/SANSKRIT/fr/>. <inria-00207929>

**HAL Id: inria-00207929**

**<https://hal.inria.fr/inria-00207929>**

Submitted on 18 Jan 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# PHONOLOGICAL OVERGENERATION IN PANINIAN SYSTEM

**Malhar Kulkarni**

IIT, Powai, Mumbai- 400076

and

M.M.Vasudevashastri Abhyankarashastri Pathashala, Pune.

malhar@iitb.ac.in

## ABSTRACT

In this paper an attempt is made to study the problem of overgeneration that is caused by the application of the system of *Pāṇini*. The system of *Pāṇini* is made up of certain rules stated by him and his commentators namely, *Kātyāyana* and *Patañjali*. These rules are supposed to produce the forms that are used in the language, i.e. Sanskrit. However, sometimes the technical application of these rules produces such forms which are not actually used in the language. In fact, sometimes it is beyond human capacities to use such forms. In the present paper two such cases dealing with the phonological overgeneration are studied and possible solutions are proposed to avoid the problem.

## 1. INTRODUCTION:

It has been demonstrated by Kiparsky and Staal(1988) how Paninian system functions on four levels, namely, semantic, deep structure, surface and phonological. This system however sometimes over-generates in perhaps, some of these levels. Of course *Pāṇini* (P) has no doubt laid down certain constraints with the help of which the system produces supposedly un-over-generated forms. Prince and Smolensky (2002), have devoted a section on Panini's theorem of constraint ranking (5.3) Of course our judgement regarding the over-generativeness of a rule in the *Aṣṭādhyāyī*(A), it must be admitted here, is based entirely upon whatever evidence in the form of pre-paninian literature available to us.

## 2. PHONOLOGICAL OVER-GENERATION

This paper is devoted to phonological over-generation that still happens with all the possible constraints applying. There are two aspects that are studied in this paper,

(1) Nasalization and (2) Phonetic doubling

### 2.1. Nasalization:

8.4.45 states that *yar*<sup>1</sup> occurring at the end of a *pada*, is optionally, (*preferably*, according to Kiparsky1980:1) substituted by the nasal, if a nasal follows.

(1)

*etad murāriḥ*  
= *etan murāriḥ* / *etad murāriḥ* ... 8.4.45

*Kātyāyana*(K) has added a *Vārttika*(V) on this rule, to the effect that this nasalization takes place permanently if the following nasal is a part of a suffix

(2)

*tad + maya*  
= *tan-maya* ... 8.4.45 + K's V.  
= *tanmaya*

#### 2.1.1. Environment for nasalization:

However, if we look at the way P has stated this rule, we have to take into account following table which shows clearly all possible environments in which this rule should apply and the possible results in the form of substitution of a nasal consonant. The top row and the left column, in the table, show the possible environment. The bottom row shows the resultant nasal consonant in place of the consonant written in the same column in top row. Thus for

<sup>1</sup>These phonemes are- all the stops including the nasals, semi vowels(y, r,l,v) and sibilants except h.

instance,

$$\begin{array}{l} [..y] \quad + [n..] / [m..] / \quad \dots \quad 8.4.45 + K's V \\ = [..y\#] \quad + [n..] / [m..] / \end{array}$$

Table 1 shows that any consonant mentioned in the top row occurring at the end of a *pada* and followed by any of the nasal consonants mentioned in the left hand column, is substituted by the nasal consonant shown in the bottom row. # mark is used to show the nasal feature in the bottom row. \* shows that these substitutions are not attested in Sanskrit. The order of sounds followed by P in his *pratyāhāra sūtras* is maintained here.

There are certain sounds in this table which are directly not applicable for this operation as they never occur at the end of a *pada* in Sanskrit. Such sounds are- *y,l,, ì , ñ , jh, bh, gh, dh,dh, kh, ph, ch*. Some grammatical entries do end in some of these sounds and hence it can be argued that by applying operations related to 0 suffix, one can generate *padas* with these sounds at the end. However, this argument does not hold valid as in the case of these consonants, the other rules namely, 8.2.30, 8.2.39 etc. will substitute them with the other consonants.

Thus consider the following example-

(3)			
<i>gumph</i>		...	<i>Dhātupāṭha</i> 6.31
<i>gumph</i>	+ <i>kvip</i>	...	3.2.178
<i>gu ph</i>	+ <i>kvip</i>	...	6.4.24
<i>gu ph</i>	+ 0	...	6.1.67
= <i>guph</i>			
<hr/>			
<i>guph</i>	+ <i>su</i>	...	4.1.1,2
<i>guph</i>	+ 0	...	6.1.68
<i>gub</i>		...	8.2.39
<i>gub/gup</i>		...	8.4.56
= <i>gub / gup</i>			

In the same way, other consonants will be substituted.

### 2.1.2. Overgenerated nasalization:

Now the rule, applied to all the remaining consonants should also apply to the following example-

(4)			
<i>catur mukha</i>	.....		8.4.45
<i>catu ṅ mukha</i>			
= <i>catu ṅmukha</i>			

However, this resultant form is not acceptable in Sanskrit. This is clearly an over-eneration.

8.4.58 states the substitution of a nasal in place of an *anusvāra* when followed by almost same consonants (called as *yay* by P) mentioned in the top row of

Table 1 above except the last three. The rule can be shown as-

$$\begin{array}{l} [...anusvāra] \quad + [yay...] \\ = [... nasal] \quad + [yay...] \end{array}$$

Thus by applying this rule we get forms like *kaṅṭha*, *aikita*, *gumṅhita* etc. Consider however, the following example-

(5)			
<i>kuṅḍam rathena</i>			
<i>kuṅḍam rathena</i>	...		8.3.23
<i>kuṅḍaṅ rathena</i>	...		8.4.58

The resultant form here is not acceptable in Sanskrit. This is again over-generation.

One may argue about redundancy being the feature of use of the *pratyāhāras* in the metalanguage of However, the tradition has taken pains in creating a constraint to check such forms in the form of statements in this regard. Pa in the context of the above example says-

*rephoṣmaṅām savarṅā na santi*<sup>2</sup>. (the sounds *r* and the sibilants do not have any homogenous(nasal))

There are at least some constraints in the form of statements of the later commentators to check the over-generation as shown above. However, in the case of phonetic doubling mentioned below, we see hardly any constraint to check the overgeneration.

## 2.2. Phonetic doubling:

P in his A has dealt with the process of reduplication at three places; (i) 6.1.1-12<sup>3</sup>, (ii) 8.1.1-15, (iii) 8.4.46-52. (i) deals with the reduplication of verbal roots in the forms of present as well as perfect tense and also in forming complex verbal roots such as desiderative and frequentative. In a nutshell, this reduplication applies to the *aiga* in Paninian terminology. (ii) deals with the reduplication of the entire *pada*. The last section in the A mentioned above, deals with the reduplication of the consonants. The paninian tradition has augmented the existing set of rules laid down by P in this section, in the form of *Vārttikas* (mainly written by K) in this regard and the later tradition has interpreted certain statements of *Patañjali*(Pa) in such a manner that the resultant forms can only be termed as over-generated ones. The later paninian tradition has done this exercise at many places and has come up with such overgenerated

<sup>2</sup>VyākaraṅaMahābhāṣya of *Patañjali*, 2001, Vol.1, p 130.

<sup>3</sup>More recently, Kiparsky in a forthcoming article available on his webpage, has discussed it.

Table 1: Consonants and their substitutes according to 8.4.45

	y	v	r	l	ñ	m	ñ	ṇ	n	jh	bh	gh	ḍh	dh	j	b	g	ḍ	d	kh	ph	ch	ṭh	th	c	ṭ	t	k	p	ś	ṣ	s	
ñ																																	
ñ																																	
ñ																																	
m																																	
	y	v	ñ	l	ñ	m	ñ	ṇ	n	ñ	m	ñ	ṇ	n	ñ	m	ñ	ṇ	n	ñ	m	ñ	ṇ	n	ñ	ṇ	n	ñ	m	ñ	ṇ	n	
	#	#	*	#																											*	*	*

Table 2: Environment for Phonetic doubling

1	2	3 Consonant Reduplicated	4	Rule of Panini
vowel	r/h	yar	—	8.4.46
—	vowel	yar	No vowel	8.4.47
vowel	Yaṇ	may	—	K & Pat on 8.4.47
vowel	may	yaṇ	—	As above
—	Śar	khay	—	As above
—	khay	Śar	—	As above

forms. The such extreme cases are presented in this paper and an attempt is made to study the approach of the Paninian system to handle this phenomenon.

(6) *putrādīnī tvam asi pāpe*

(Oh! son-eater woman, shame on you!)

*putrādīnī sarpiṇī*

(she-snake is son-eater. )

In this case, t is seen reduplicated alongwith the change in the meaning. This case is noted by 8.4.48.

### 2.2.1. Environment for Phonetic doubling:

In the same section, some other phonemes are also noted for their reduplicated occurrence. K and pat have also noted down this tendency in some other phonemes. These phonemes are- same mentioned in fn 2. In table 2 they are referred to as yar, as used by P. In the table 2, these rules are explained with all details, namely environments- prior and posterior

Here 1 , 2 , 4 refer to the environment for phonetic doubling. The order indicates the positions of these environments and the position of the phoneme reduplicated. The examples for these two rows are-

(7) *haryyanubhavaḥ*

(h ā-r-y anubhavaḥ > phonetic doubling of y)

(8) (a) *rāmātt*

(rām ā-t-(no vowel) > phonetic doubling of t)

(b) *sudhdhyupāsyaḥ*

(s-u-dh-y upasyah > phonetic doubling of y).

### 2.2.2. K and Pa on the environment for phonetic doubling:

While commenting upon 8.4.47, K notes- *dvirvacane yaṇo mayaḥ* . On this Pa has a two fold comment. He says-

*dvirvacane yaṇo maya iti vaktavyam.*

*Kim udāharaṇam yadi yaṇa iti*

*pañcamī maya iti śaṣṭhī*

*ulkkā valmmikam ity udahāraṇam. Atha maya*

*iti pañcamī yaṇa iti śaṣṭhī*

*dadhyatra madhvvatreyi udāharaṇam*

This means- In the rules dealing with the process of phonetic doubling, the words *yaṇo mayaḥ* should be stated. What is the example ? If *yaṇaḥ* (*yan* is y, v,r, l) is taken to be ablative and *mayaḥ* (*may* is all stops except nasal palatal) is taken to be genitive, then the examples are –

(9) *ulkkā / valmmikam*

and if *mayaḥ* is taken to be ablative and *yaṇaḥ* is taken to be genitive, then the examples are-

(10) *dadhyatra / madhvvatraya*.

Same argument is applied to another statement of K, namely *ś arahḥ khayāḥ*<sup>4</sup> which provides us with the following examples-

(11) *sththālī / sththātā*

(12) *vatssaḥ / kssīram / apssarāḥ*

This way of interpreting the statements of K on the rules of P becomes a peculiar feature of the system of paninian grammar. Later tradition of paninian grammar thus by interpreting statements of K and Pa and

<sup>4</sup>This statement means that *khay* is reduplicated if it occurs after śar and śar is reduplicated if it occurs after *khay*. śar stands for all the sibilants except h and *khay* stands for all the voiceless stops.

P have noted down forms which we here address as overgenerated forms.

We note that this feature is also noted by non-paninian systems such as *Kātantra*. A commentary on *PrakriyāKaumudī* namely *Prakāśa* notes that according to *Kātantra* school the phonetic doubling in a particular case will give rise to only 32 forms and not more<sup>5</sup>.

### 2.2.3. Twice Occurrences of same consonant in Sanskrit

It is noteworthy to study the structure of the consonant cluster in Sanskrit. A list of such clusters is available in Coulson Michael, 2003, p 22-24. We concentrate on a cluster of two consonant of same phonetic value. In other words, we concentrate on the twice occurrence of the same consonant. In the table 3, a list of such consonant clusters is provided. Table 3 shows us the consonants which can have twice occurrence without applying the rules of phonetic doubling.

Table 3: Twice occurrences of same consonant

<i>k</i>	(i)Final + initial of the next word (ii)Prefinal
<i>g</i>	Final + initial of the next word
<i>c</i>	Final + initial of the next word
<i>j</i>	Final + initial of the next word
<i>t</i>	Final + initial of the next word
<i>d</i>	Final + initial of the next word
<i>p</i>	Final + initial of the next word
<i>b</i>	Final + initial of the next word
<i>ṅ</i>	Final + initial of the next word
<i>n</i>	(i) Final + initial of the next word (ii)Pre-final
<i>ṇ</i>	Final + initial of the next word
<i>m</i>	Final + initial of the next word
<i>ś</i>	Final + initial of the next word
<i>ṣ</i>	Final + initial of the next word
<i>s</i>	Final + initial of the next word

A careful glance at table 3 will point out that all these consonants fall in the domain of the application of the phonetic doubling rules mentioned above in Table 2. Therefore, if the rule for phonetic doubling is applied to these already existing two consonants, we get three same consonants occurring one after another. Such a form is noted to exist optionally by P in the case of consonants except nasals by the 8.4.65.

<sup>5</sup>*PrakriyāKaumudī*, 2000, Vol.I, p 158.

### 2.2.4. Generation of Phonetic doubling in later tradition:

A 17<sup>th</sup> century grammar text, *Vaiyākaraṇa-SiddhāntaKaumudī* (VSK) records following cases of phonetic doubling-

(13) *rāmātt rāmādd* / VSK 206, *dvitve rūpacatuṣṭayam*. (in the forms *rāmāt* and *rāmād*, after applying the rules of phonetic doubling we get 4 forms).

(14) *aidhidhvam* / VSK 2258, *dhaḍhayor vasya masya ca dvitvavikalpāṣoḍaṣarūpāṇi*. (by reduplicating *v* and *m* when immediately before *dha* and *ḍha* we get 16 forms.)

(15) *saṃskartā* / VSK 138, *anusvāravatām anusvārasyāpi dvitve dvādasa*. (after reduplicating the *anusvāra* in the forms already containing it, we get 12 forms).

(16) *gavāk* / VSK 443

Cases (15) and (16) deserve a special attention as they pose a problem.

#### 2.2.4.1. 2.2.4.1 Generation of Phonetic doubling in the forms of *saṃskartā*

(15) *saṃskartā* - This word is formed in the following way<sup>6</sup>-

*sam + kartā*  
*sam + s- kartā* ... 6.1.134  
*sar + s-kartā* ... 8.3.5  
*saṃr + s-kartā* ... 8.3.2 / 8.3.4  
*saṃs + s-kartā* ... 8.3.15

Along with this form there is an optional form that is available in which in place of *m* there occurs an *anusvāra*. In the following two tables (Table 4 and Table 5), forms with *m* and *anusvāra* are presented.

In Table 4 and 5, we see phonetic doublings of *s*, *t*, *k* and more problematically of the *anusvāra*. This phonetic doubling of *anusvāra* is based on the argument of K that *ayogavāha*<sup>7</sup> *s* are to be included in the *pratyāhāra aṭ* as well as *śar* by the statement-*ayogavāhānam aṭṣuṇatvam śarṣu jaś tvaṣatve*.

<sup>6</sup>I have to turn to Devnagari fonts for these two case to stress the amount of problem.

<sup>7</sup>The term *ayogavāha* refers to *anusvāra*, *visarga*, *jihvāmūliya* and *upadhmanīya*, VyākaraṇaMahābhāṣya of Patañjali, 2001, Vol.1, p 132.



The reasoning for the phonetic doubling of *k*, *g*, *ñ* is 8.4.47. The reasoning for the phonetic doubling of *y* and *m* is the same as mentioned in Table 5 namely, *dvirvacane yaṇo mayah*. We also note that there is a phonetic doubling of even a *visarga* in certain forms. The reasoning for this phonetic doubling is same as mentioned after Table 5, namely, inclusion of *visarga* in the *pratyāhāra yar*. Also there is nasalization which is marked by a sign on certain forms which has added those many forms.

We note that in the table 11 (as shown in the Appendix) the following consonants apart from the ones mentioned in Table 7 are reduplicated- *ṛ*, *ṣ*. The reasoning for phonetic doubling of *ṛ* is 8.4.47 and for *ṣ* is the one mentioned in Table 4. namely, *ś arah khayah*. We also note that in some forms even the *visarga* is reduplicated like in the previous table. In Table 10 and 11 (as shown in the Appendix), we also note that in some forms three consonants are simultaneously reduplicated. We also see that nasalization is marked with the sign in some forms.

Thus if we compare the tables 10 and 11 (as shown in the Appendix) statistically we come up with the following picture-

Unduplicated	Duplicated + Nasalized
49 (Table 8)	196 (Appendix: Table 10)
69 (Table 9)	267 (Appendix: Table 11)

If we are adopting the Paninian framework for generating forms by machine we will face similar problems if we apply the rules of phonetic doubling .

### 3. PROPOSED SOLUTION :

This overgeneration of forms is caused by-

- (i) redundancy of the *pratyāhāra*.
- (ii) application of the rules of phonetic doubling mechanically.
- (iii) application of statements and interpretations of later paninian commentators.

To solve this problem we propose the following:

If we are going to apply the rules of phonetic doubling we must make a rule that -

R1 The visarga should never be reduplicated.

R2 An anusvara should never be reduplicated.

R3 The rule of phonetic doubling should not be applied more than once to one consonant.

R4 The rule of phonetic doubling should not be applied to more than one consonant simultaneously.

In order to remove the redundancy, we have to rely upon the statements of the later comentators and take note of their statements and modify the rule accordingly.

### 4. ACKNOWLEDGEMENT

I wish to express gratitude to my students Ms.A.Ajotikar, Ms.T.Ajotikar and Ms.Sarnaik, at the Abhyankarashastri Pathashala, Pune for helping me type out the forms in tables presented in this paper. I also wish to express my gratitude to all the scholars who made suggestions and remarks which helped improve this article immensely. I wish to thank Prof. Kiparsky for providing me with the details of the reference of one of his forthcoming publications. I wish to thank also my student Ms. Chatali Dangarikar for helping me format the text of this article.

### 5. REFERENCES

K. V. Abhyankar, editor. 1943. *Vyākaraṇa-Mahābhāṣya of Patañjali with Marathi Translation (7volumes)*. D.E.Society. reprint, Sanskrit Vidya Parisamstha, Pune, 2007.

Prof. Balshastri, editor. 2001. *Vyākaraṇa-Mahābhāṣya of Patañjali, alongwith Pradīpa, Udyota and Śabdakaustubha*,. Pratibha Prakashan, New Delhi. edited originally by Shri. GuruprasadShastri, reedited by Prof. Balshastri.

B.K.Dalai, editor. 2007. *Phonology, in Studies in Sanskrit Linguistics*,. Bharatiya kala Prakashan, New Delhi.

Otto Böhtlingk, editor. 1998. *Pāṇinis Grammatik*. Motilal Benarsidass, New DELhi, 1<sup>st</sup> indian edition edition. Aṣṭādhyāyī edited and translated into German.

George Cardona and Dhanesh Jain, editors. 2003. *Indo-Aryan Languages*. Number 11 in Routledge Language Family Series. New Fetter Lane, London EC4P4EE.

George Cardona. 2004. *Recent researches in Paninian studies*. Motilal Benarsidas, , Delhi, 2nd revised edition edition.

Michael Coulson. 2003. *Teach yourself Sanskrit*. Hodder Headline Ltd., revised by prof.richard gombrich and dr james benson edition. 1st published in 1976.

Dr. GirijaKumar Dixit, editor. 1987. *Paribhāśenduśekhara of Nāgeśa, alongwith the commentary Sarvamāṅgalā*,. Sampurnananda Sanskrit University, varanasi edition.

Dr. Sitaramashastri, editor. 1996. *Bṛhacchabdenduśekhara of Nāgeśa (in 3 parts)*. Sampurnananda Sanskrit University, Varanasi, 1<sup>st</sup> edition edition.

G.Caturveda and P.Bhaskar, editors. 1987. *Vaiyākaraṇa-Siddhānta-Kaumudī, alongwith Bāla-Manoramā and Tattvabodhinī*,. Motilal Benarsidass.

Paul Kiparsky and Staal J.F, 1988. *Modern Studies in Sanskrit*, chapter Syntactic and Semantic relations in Panini.

Paul Kiparsky. 1980. *Panini as a variationist*. Centre of Advanced Study in Sanskrit,University of Poona, in collaboration with MIT press,, Cambridge(Massachusetts, U.S.A.) and London (England).

Paul Kiparsky. (forthcoming). Reduplication in stratal ot. Available at: <http://www.stanford.edu/~kiparsky/Papers/reduplication.pdf>.

Pt. Muralidhar Mishra, editor. 2000. *Prakriyākaumudī with Prakāśa (in 3 parts)*. Sampurnanada Sanskrit University, Varanasi.

## Abbreviations

A	-	<i>Aṣṭādhyāyī</i>
K	-	<i>Kātyāyana</i>
P	-	<i>Pāṇini</i>
Pa	-	<i>Patañjali</i>
V	-	<i>Vārttika</i>
VSK	-	<i>Vaiyākaraṇa-Siddhānta-Kaumudī</i>



Appendix: Table 10:Phonetic doubling in the declension of *Gavāc*

1	2	3
1 गोऽक्, गोऽग् गोअक्, गोअग् रस्रक्, गवऱग् 6	गोची	गोऽङ्चि गोऽङ्चि गोऽङ्चि गोअङ्चि गोअङ्चि गोअङ्चि गवऱङ्चि गवऱङ्चि गवऱङ्चि 9
2 गोऽक्, गोऽग् गोअक्, गोअग् रस्रक्, गवऱग् 6	गोची	गोऽङ्चि गोऽङ्चि गोऽङ्चि गोअङ्चि गोअङ्चि गोअङ्चि गवऱङ्चि गवऱङ्चि गवऱङ्चि 9
3 गोची	गोऽग्भ्याम् गोऽग्भ्याम् गोऽग्भ्याम् गोअग्भ्याम् गोअग्भ्याम् गोअग्भ्याम् रस्रग्भ्याम् रस्रग्भ्याम् रस्रग्भ्याम् गोऽग्भ्याम् गोऽग्भ्याम् गोऽग्भ्याम् गोअग्भ्याम् गोअग्भ्याम् गोअग्भ्याम् रस्रग्भ्याम् गवऱग्भ्याम् गवऱग्भ्याम् गोऽग्भ्याम् गोऽग्भ्याम् गोअग्भ्याम् गोअग्भ्याम् रस्रग्भ्याम् रस्रग्भ्याम् 24	गोऽग्भिः गोऽग्भिः गोऽग्भिः गोऽग्भिः गोअग्भिः गोअग्भिः गोअग्भिः गोअग्भिः गवऱग्भिः रस्रग्भिः गवऱग्भिः रस्रग्भिः 12
4 गोचे	Same as above 24	गोऽग्भ्यः गोऽग्भ्यः गोऽग्भ्यः गोऽग्भ्यः गोअग्भ्यः गोअग्भ्यः गोअग्भ्यः गोअग्भ्यः गवऱग्भ्यः रस्रग्भ्यः रस्रग्भ्यः रस्रग्भ्यः गोऽग्भ्यः गोऽग्भ्यः गोऽग्भ्यः गोऽग्भ्यः गोअग्भ्यः गोअग्भ्यः गोअग्भ्यः गोअग्भ्यः गवऱग्भ्यः रस्रग्भ्यः रस्रग्भ्यः रस्रग्भ्यः 24
5 गोचः	Same as above (24)	Same as above 24
6 गोचः	गोचोः	गोचाप्
7 गोचि	गोचोः	गोऽङ्चि गोऽङ्चि गोऽङ्चि गोऽङ्चि गोऽङ्चि गोअङ्चि गोअङ्चि गोअङ्चि गोअङ्चि गोअङ्चि गवऱङ्चि रस्रङ्चि गवऱङ्चि रस्रङ्चि गवऱङ्चि रस्रङ्चि गोऽङ्चि गोअङ्चि गवऱङ्चि 24

