Integrity Of Geminates And Homorganic Nasal Plus Stop: Evidence From A Word Game

Manjari Ohala
San José State University
ohala@socrates.berkeley.edu

ABSTRACT

Geminates and homorganic nasal + stop (NC) clusters have been the topics of speculation about their possessing or lacking 'integrity' (unitary character). I present experimental evidence on this issue as it pertains to such entities in Hindi. Evidence from a new experimental technique, a word game, shows that geminates are treated as indivisible but for NC clusters the situation is more complex: bilabial and alveolar NCs were treated by subjects as divisible but velar NCs were largely kept intact. The results for the retroflex and palatal NCs were inconclusive. A voiced stop in NC clusters significantly influenced responses that did not break up the NC.

1. Introduction

Intervocalic geminates have traditionally been considered clusters and syllabified VC-CV [4][5]. Recent proposals in moraic phonology [3] [11] discuss the notion of the 'integrity' of geminates, e.g., they are not subject to rules of epenthesis, and treat them as constituting the moraic coda of one syllable and the non-moraic onset of the next. Intervocalic homorganic NC sequences have, with few exceptions (e.g. Luganda), been treated as clusters of N + stop. I present experimental data from Hindi supporting the 'integrity' of geminates but showing that the NC sequences present a more complex picture in that they do not uniformly form a cohesive group.

At the phonetic level Hindi has the following nasals [m], [n], [ɳ], [ɲ], [ŋ]. Based on the usual criteria of contrast and complementary distribution /m/ and /n/ would be judged phonemes. The nasals [ɳ] and [ɲ] only occur before homorganic consonants. The status of [ŋ] is somewhat controversial in that it occurs before non-velar segments only in a few Sanskrit loanwords such as [vaŋməj] 'literature', otherwise it occurs only before homorganic stops. Based on this some linguists consider it a phoneme and others an allophone of /n/. The only time [ɳ], [ɲ], or [ŋ] occur independently is in reciting Devanagari, which, following the Sanskrit orthography, has separate symbols for all five nasals. Thus in recitation one would say [kə k^hə gə $g^ɦ$ə ŋə] and similarly for the other places of articulation. Also, it should be mentioned that the symbols for the retroflex, palatal, and velar nasals only occur in the spelling of a few Sanskrit loans. Homorganic nasals are written either with a dot over the vowel or with the symbol for a half consonant attached to the next consonant.

Phonetically, geminates involve the closure held for a longer period vis-à-vis their non-geminate counterparts. Thus phonetically they are not really two separate consonants as the term "geminate" implies. The Devanagari orthography does however represent them as two consonants (using a half symbol for the first consonant and thus overtly indicating a cluster). Almost all of the consonants of Hindi can occur as geminates [8] and they are phonemic , i.e. they are "true" geminates in current phonological terms [11]. However, they have rather severe distributional restrictions: they only occur intervocalically and the preceding vowel must be short. Such restrictions are not shared by the other single consonants nor by any consonant clusters, both of which—including -NC- clusters—can be preceded by long vowels.

2. Experimental Evidence

The results from earlier experiments I conducted using the Fallows [2] and Derwing [1] experimental techniques for –NC- and geminates were inconclusive because of problems with pronunciation constraints (for details see [9]). Thus it was desirable to develop another experimental technique which would help to access mental reality by circumventing the problems associated with phonotactic/pronunciation constraints. For words involving geminates, it is also not clear how orthographic influence would manifest itself. On the one hand since geminates are always written as clusters, one could make a case for the influence to be in the direction of VC-CV. On the other hand since the first consonant of the geminate (i.e. the first part of it) is written with the symbol for a half consonant and such a symbol only occurs attached to another consonant, one could also make a case for a VCC-V or V-CCV influence (in either case violating the phonotactic constraints mentioned above).

A new experimental technique: In order to circumvent the problem referred to above of pronunciation constraints impacting on the experimental results, I developed a variation on the concept formation technique [6] and used it to address syllabification of -NC- and geminate.

ISBN 1-876346-48-5 © 2003 UAB

Experiment A

Method: Subjects were told that a new game for children was being devised in which [ɪ] is inserted in the middle of a word. Then I gave two examples. In the case of intervocalic single consonants they were trained to insert it before the 'C': [t͡ʃʰʊpa] 'hidden' → [t͡ʃʰʊɪpa]. For intervocalic clusters, [ɪ] was inserted between the 'CC': [bəsta] 'knapsack' → [bəsɪta]. I then continued with the "training" session, giving them feedback orally: the correct answer if they made a mistake; and saying 'correct' if their answer was correct. During this 'training' session there were 45 words. I stopped once subjects reached criterion at 13 right in a row. At the end of the training session I went on to the "test" session. Subjects were not told when the training session ended and the test session started. However, in the test session no feedback was given.

Word lists: The 45 item word list used in the training session consisted of 22 words with intervocalic single consonants and 23 with a variety of intervocalic two-consonant clusters. However none involved NC or geminates. The "test session" consisted of 11 words: 2 with intervocalic single consonants, 3 with intervocalic clusters similar to the training session. These 5 words also acted as a control to see if the subjects had indeed learned the word game since they were similar to the words used in the training session. (If subjects made mistakes on these it would be an indication that they hadn't learned the word game.) There were four words with intervocalic geminates (given in Table 7) and 2 with intervocalic NC clusters. Of the 4 words with geminates, two contained stops (one voiceless one voiced), one a fricative, and one a nasal. Thus these geminates represented the spectrum on the sonority scale. The NC clusters were [-ŋkʰ-] and [-ɳɖ-] (the words used were [pəŋkʰa] 'fan' [ɖəɳɖa] 'stick'). The vowel [ɪ] was chosen as the vowel for insertion because geminates are only preceded by short vowels. So the choice would have been from [ɪ ʊ ə]. [ə] and [ʊ]were not chosen because many eligible words in the language already had these vowels in them and thus their insertion could make for awkward pronunciation.

Subjects: The subjects were 21 native speakers of Standard Hindi residing in New Delhi, India. They were all literate and all of them also spoke English. The experiment was administered to them individually by me. Three of the subjects did not reach criterion in the training session and three did not give correct answers to the control words in the test session. Thus the results to be presented are based on 15 subjects.

Results And Discussion: The results for the geminates are given in Table 1. Here –ɪC:- indicates a geminate kept as a unit (i.e. [ɪ] was inserted before the geminate). -CɪC- indicates [ɪ] was inserted in the middle of the geminate (i.e. the geminate was treated as a cluster). The responses listed under 'other' were either of the type where the vowel of the first syllable was changed, for example, in response to [gəd:a] if the subject responded [gɪd:a], or in responses to [gən:a] they gave [gɪn:a]. Or they were of the type where [ɪ] was added at the end of the word as in [gədəɪ] for [gəd:a] etc. Here and elsewhere the responses in the 'other' column have not been included in the statistics since they represent cases where subjects were not following the pattern taught in the training session.

Table 1. Results for the syllabification of intervocalic geminates.

Word ↓ Response Type →	-ɪC:-	-CɪC-	other
[pək:a] 'firm'	9	5	1
[gəd:a] 'mattress'	9	4	2
[gən:a] 'sugarcane'	9	3	3
[gʊs:a] 'anger'	9	4	2
Totals	**36**	**16**	**8**

Ignoring the 'other' responses, the majority responses kept the geminate intact (χ^2 = 7.692308, df = 1 $p < .01$) Here orthography was not an influencing factor since, as mentioned above, it would have favored the cluster responses given that geminates are always written as clusters in Devanagari. Given these results, the syllabification VC-CV would not be justified. The results of this experiment seem to indicate that in the mental grammar of these native speakers the integrity of geminates is maintained, i.e. they are treated as units, not clusters. This of course does not preclude their duration as being more than that of singletons, i.e., having more weight than singletons on the timing tier as is claimed by current phonological theory.

Although the results for –NC- showed there was a statistically significant tendency toward keeping these clusters intact (χ^2 = 4.764706, df =1, $p < .05$), these results were not clear cut. In the case of the velar nasal they seem to point to the cluster being kept intact and thus perhaps the syllabification VNC-V. (V-NCV would not work because it would violate phonotactic constraints.) However, the retroflex nasal did not show such a trend , and in fact it elicited a large number of 'other' responses. Given that only two NC clusters were included in this experiment with one token each, it was desirable to conduct another experiment with a wider variety of NC clusters as well as non-homorganic nasal plus stop clusters. The next experiment was conducted with this aim in mind.

Experiment B

Procedure: The same experimental technique described above was used except this time in the training session I went

through all 45 words even if the subjects had already got 13 right in a row so that all subjects would have equal chance at hearing all the words and the correct answer to any incorrect responses.

Word list: The word list used in the training session was the same except seven words were replaced by others. (These seven had given a few of the subjects of the previous experiment some problems in ɪ-insertion for reasons I cannot explain.) This time the word list for the test session contained 25 items and all of the words of interest were with nasals. There were 10 control items (i.e. words similar to the training session), 4 with intervocalic single consonants and 6 with clusters. The NC clusters included the velar, retroflex, palatal, bilabial, and alveolar nasals followed by homorganic stops (both voiced and voiceless). In addition there were 4 words with *nasal + non-homorganic stop* clusters: two words had the bilabial nasal followed by a velar or alveolar stop and two words had the alveolar nasal followed by a velar or bilabial stop. (Table 2 lists all of the words with *nasal + stop* clusters.)

Subjects: The subjects were 21 native speakers of Standard Hindi residing in Delhi (not the same as those in experiment A). They were either high school or college students except one subject who was semiliterate—he was an adult who had not gone beyond 2nd grade and was unable to read Devanagari with any proficiency. His results did not differ in any significant way from the other subjects, and so I have included them. Six did not learn the word game therefore the results are based on 15 subjects.

RESULTS AND DISCUSSION

The results are presented in Table 2. '-ɪNC-' indicates the *nasal + stop* cluster was kept intact. -NɪC- = the nasal was pronounced as such (i.e., velar as velar, palatal as palatal, etc) followed by the -ɪC-. -nɪC- = the place of articulation was changed to an alveolar/dental nasal in the pronunciation. 'Other' = responses where changes such as the following occurred: changing the vowel ([bəndɪr] for [bəndər]) or if [ɪ] was inserted after the cluster ([tambɪa]) etc. In doing the statistics, the 'other' responses have not been included.

It is clear from looking at the responses to the last four words that for non-homorganic nasal plus stop clusters a VC-CV syllabification would be justified since [ɪ] was overwhelmingly inserted between them (χ^2 = 102.9, df = 3, p <.001). The orthography also writes the nasal here with a symbol for a full nasal (i.e. as [n] or [m]). In the case of the bilabial and alveolar *homorganic nasal + stop* clusters also a VC-CV syllabification is justified because the majority of the subjects did break up the cluster with the [ɪ] (respectively, χ^2= 15.9, df = 2, p < .001, χ^2 = 31.7, df = 2, p < .001). In these cases when the cluster was broken up, the nasal's place of articulation was not changed. As mentioned above [n] and [m] are the two nasals that occur freely, i.e. are clearly phonemic.

Table 2. Results for the syllabification of intervocalic nasal plus stop clusters.

Response→ Word ↓	-ɪNC-	-NIC-	-nɪC-	other
[pəŋkʰa] 'fan'	8	5	0	2
[guŋga] 'mute'	13	0	2	0
Totals	**21**	**5**	**2**	**2**
[gʱənʈa] 'hour'	4	0	10	1
[dəɳɖa] 'stick'	9	0	6	0
Totals	**13**	**0**	**16**	**1**
[kəɲt͡ʃən] 'gold'	2	0	11	2
[bʱaɲd͡ʒa] 'nephew'	7	0	7	1
Totals	**9**	**0**	**18**	**3**
[t͡ʃampa] '*micalea campaca*'	2	12	0	0
[tamba] 'copper'	9	5	0	1
Totals	**11**	**17**	**0**	**1**
[əntər] 'between'	1	12	0	1
[bəndər] 'monkey'	3	10	0	12
Totals	**4**	**22**	**0**	**4**
[kʊnba] 'clan'	2	13	0	0
[tɪnka] 'twig'	1	14	0	0
[nəmda] 'rug'	0	15	0	0
[d͡ʒʱumka] 'earring'	0	15	0	0
Totals	**3**	**57**	**0**	**0**
Totals: N + Vls stop	**7**	**17**	**29**	**21**
Totals: N + Vd stop	**4**	**41**	**15**	**15**

The responses to the other NC clusters are a bit more complex. In the case of velar nasals, the overwhelming response was to keep the nasal plus stop cluster intact (i.e. not insert an [ɪ] between the nasal and the stop) (χ^2 =22.4, df = 3, p < .001). Thus a VC-CV syllabification would not be justified. For the retroflex and palatal places of articulation, there was no clear preference between keeping the NC intact vs inserting the [ɪ] between the nasal and the stop but changing the place of the nasal to alveolar, presumably due to pronunciation constraints since these nasals only occur before homorganic consonants. All that can be concluded is that there was a clear preference *not* to break up the cluster while retaining the original place of articulation of the nasal (respectively, χ^2 = 15, df = 2, p < .001; χ^2 = 18, df = 2, p < .001) Finally, I also examined the difference between the NC clusters where the 'C' is voiceless vs those where it is voiced. There is a clear difference between these two cases.

ISBN 1-876346-48-5 © 2003 UAB

In *N + voiceless stop* clusters the distribution is not statistically significant ($\chi^2 = 3.34$, df = 2, p = n.s.) but in the case of *N + voiced stop* it is ($\chi^2 = 19$, df = 2, $p < .001$). The subjects made a clear choice not to break up the cluster only in the latter set. Ohala & Ohala [10] discuss the asymmetric behavior of NC clusters depending on whether the 'C' is voiced or voiceless and present a possible explanation for this asymmetry.

Orthographic influence is not at issue here since there is no discernable correlation in the treatment of the words by the subjects and their orthographic representation. All of the -NC- words at issue are written with the diacritic for homorganic nasals except [tamba] and [guŋga] which are written with the diacritic for a nasalized vowel (although they are pronounced with -NC-, cf. M. Ohala [7]) and [bʰaɲd͡ʒa] which is written with the symbol for a full alveolar nasal.

What is clear from the findings presented here is that all -NC- clusters do not all behave alike. Before assigning the syllabification VC-CV for all of them, more research needs to be done. For example, what is the duration of these nasals before voiced and voiceless consonants? If the duration of the nasal is long in comparison to that of the voiced stop, it may account for the preference not to break the cluster. This may be particularly important in the case of the *velar nasal + voiced stop* sequences. Some of the acoustic measurements I did on data unrelated to this experiment seem to show that for some words with *velar nasal + voiced velar stop* clusters the stop is entirely missing. Could this possibly account for the differential behavior of the velar nasal? Also, what is the exact place of articulation of the retroflex and palatal nasals? Could it be that the retroflex and palatal nasals are really retracted alveolars (i.e. post-alveolar)? If so, such clusters might be considered by some subjects to be non-homorganic and thus capable of being broken up. Palatographic data would be useful. All of these are topics for further research. In any future test of this type it would be useful to tape-record the subjects' responses so that acoustic analysis could also be done. Experimental studies on other languages would also be useful.

CONCLUSION

The results of the experiments presented here show that geminate consonants do have 'integrity' in that they resist division. Homorganic nasal + stop clusters, however, exhibit more complex behavior. Labial and dental clusters of this type were divided by subjects whereas velar clusters were generally kept intact. The results for retroflex and palatal clusters were inconclusive. Unexpectedly, when the stop in these clusters was voiced there was a significantly greater tendency to keep all NC clusters intact. These result also show that issues of syllabification and of the integrity/divisibility of sound sequences must not only be informed by universal phonetics but also by language-specific psycholinguistic factors.

REFERENCES

[1] Derwing, B. L. (1992a). A 'pause-break' task for eliciting syllable boundary judgments from literate and illiterate speakers: preliminary results for five diverse languages. *Language and Speech,* 35, 219-235.

[2] Fallows, D. 1981. Experimental evidence for English syllabification and syllable structure. *Linguistics*, 17, 179-392.

[3] Hayes, B. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20.253-306.

[4] Kelkar, A. R. 1968. *Studies in Hindi-Urdu I: introduction and word phonology*. Building Centenary and Silver Jubilee series, 35. Poona: Deccan College.

[5] Maddieson, I. 1985. Phonetic cues to syllabification. In: Fromkin, V. (ed) Phonetic Linguistics. Orlando: Academic Press. 203-221.

[6] Ohala, J. J. 1986. Consumer's guide to evidence in phonology. *Phonology Yearbook* 3.3 - 26.

[7] Ohala, M. (1983). *Aspects of Hindi Phonology*. Delhi: Motilal Banarsidass.

[8] Ohala, M. 1994. Hindi. *Journal of the International Phonetic Association* 24:1.35-38.

[9] Ohala, M. 1999b. Hindi Syllabification. *Proc., ICPhS 1999,pp.* 727-730.

[10] Ohala, M. & Ohala, J. J. 1991. Nasal epenthesis in Hindi. *Phonetica* 48.207-220.

[11] Perlmutter, D. (1995). Phonological quantity and multiple association. In Goldsmith J. J. (Ed.) *The handbook of Phonological Theory* (pp. 307-317). Oxford: Blackwell.