# Acoustic VC transitions correlate with degree of perceptual confusion of place contrast in Hindi

Manjari Ohala and John J. Ohala

## 1. Introduction

Speech sounds must be (a) different from their neighboring sounds (a syntagmatic requirement) and (b) different from all the other sounds that might have appeared in the same context (a paradigmatic requirement). Common cross-language sound patterns manifesting such constraints suggest that they arise from physical phonetic factors, the only thing common to all languages. For example, for a syntagmatic constraint, in English (and many other languages (Kawasaki 1982; Ohala & Kawasaki-Fukumori 1997)) in native vocabulary the labial-velar approximant [w] fails to occur as C2 in syllable initial consonant clusters where C1 is [labial]. This is presumably motivated by the fact that the consonantal transitions from a labial consonant create a formant trajectory that is too similar to that of a [w] to be reliably differentiated from it. Thus, although /w/ is otherwise paradigmatically distinct from other approximants that can appear in C2 position, i.e., /j l r/, it is syntagmatically too close to other labials. Similar acoustic-perceptual factors account for paradigmatic phonological constraints. For example, distinctive pharyngealization in Arabic is manifested more easily on coronal consonants than labial or velar (in other words, pharyngealization's distinctive lowering of F2 is not as easily detectable on segments like labials and back velars that already have a low F2 (Ohala 1985)). Thus by most analyses of Arabic, only coronals exhibit distinctive pharyngealization (Ghazeli 1977). The acoustic-perceptual constraints on speech sounds' distinctiveness impacts on a number of issues.

First, how capacious is the 'phonological space'?[1] From cross-language studies it is possible to derive a notion of the "usual" or modal number of contrasts one can expect in certain features (Maddieson 1984). For example, vowel height typically has three contrasts and VOT typically has two. Fewer or more contrasts than these are possible but if more than the modal number are used one expects some "cost", e.g., in the production domain where tighter control of relevant phonetic parameters may be required (thus requiring no extra cost at the perceptual level) or in the perceptual domain where the cost may be more confusion between the relevant sounds (and thus implying no extra cost at the level of production). (See also Zipf 1935; Lindblom & Maddieson 1988; Lindblom 1984, 1989.)

Second, the fact that the set of multiple acoustic cues differentiating most features may be different in different environments would predict that there should be context-specific constraints on features. For example, stop bursts contain important cues for their place of articulation (Fischer-Jorgensen 1954) but for aerodynamic reasons they only have bursts upon their release, not on their onset. The presence of these cues only upon release may explain in part why CV syllables are more common than VC (J. Ohala 1995) and why stop place is often neutralized in coda position or why medial heterorganic stop clusters yield a stop (geminate or singleton) that has the place of the C2, the stop that had the release, e.g., Late Latin /nokto/ > Italian /not:e/ 'night' (Ohala 1990). Ohala (1975) and Ohala and Ohala (1993) speculated that the reason why there are often more place contrasts among nasals in post- rather than pre-vocalic position (contrary to the pattern for obstruents) is because nasals have greater velopharyngeal opening post-than pre-vocalically, and since the strength of the spectral differences differentiating nasals by place depend on the degree of nasal-oral coupling. (See also Ohala & Kawasaki 1984, Ohala 1992, Ohala & Kawasaki-Fukumori 1997.)

One might suppose that paradigmatic and syntagmatic similarity would especially be a problem in languages that supported a relatively large number of contrasts. Hindi, for example, has 34 consonant distinctions, not counting singleton (short) vs. geminate (long) contrasts. Among stops (including affricates) it has four laryngeal distinctions (voiced, voiceless, voiceless aspirated, and breathy-voiced) and five



place distinctions (labial, dental, retroflex, palatal, velar) (M. Ohala 1994). Table 1 lists the segments of Hindi and Table 2 presents a minimal set showing the five place contrasts in coda position.

| | | | | | |
|---|---|---|---|---|---|
| p | t | ʈ | t͡ʃ | k | |
| pʰ | tʰ | ʈʰ | t͡ʃʰ | kʰ | |
| b | d | ɖ | dʒ | g | |
| b̤ | d̤ | ɖ̤ | d͡ʒ̤ | g̈ | |
| f | s | | ʃ | | |
| | z | | | | |
| m | n | | | ŋ | |
| ʋ | | | j | | h |
| | ɾ | ɽɽ̈ | | | |
| | l | | | | |

Table 1. Consonant inventory of Hindi (excluding geminates).

| | |
|---|---|
| pip | *pus* |
| pit̪ | *yellow* |
| piʈ | *beat (verb)* |
| pit͡ʃ | *raised garden row* |
| pik | *spit (betel juice)* |

Table 2. Hindi words (and their English translations) showing contrast of five places of articulation in stops in coda position.

Hindi thus presents opportunities to explore this issue along with a related issue brought up by a recent claim of Steriade (1997) regarding the context where retroflex place distinction is most robust. Steriade starts with the view, similar to that of Ohala & Kawasaki (1984) and Ohala & Kawasaki-Fukumori (1997), that contrasts are "...permitted (or licensed) in positions that are high on a scale of perceptibility."

Most of her arguments are based on neutralization of certain laryngeal contrasts but, referring to a study by Dave (1977), she also considers the case of retroflexes and indicates that since the VC transitions of retroflexes are more prominent than their CV transitions, the position that should show more neutralization involving retroflexion is the onset (CV) position (i.e. word-initial or post-consonantally). (This is in contrast to other segment types where CV cues are said to be more prominent than VC.) Anderson (1997) finds support for Steriade's claims based on perceptual data from Western Arrernte intervocalic dental/alveolar/retroflex contrasts. She found that the vowel preceding the retroflex shows the most prominent formant transitions and these formant transitions increase correct identification more than is the case with alveolar stops.

The VC (rather than CV) context is, perhaps, an ideal testing ground for an investigation of the costs of accommodating more than the usual number of place contrasts because there are circumstances where the set of differentiating cues could be quite impoverished. The place of the C in a VC context has some cues at its onset, namely from the VC formant transitions, as well as from its offset, namely the release burst. However, in many cases the stops at least may be unreleased.

Thus if the 'cost' attached to this higher-than-modal number of place contrasts is paid in tighter control at the production level, one would expect to find lesser overlap of these place-cueing transitions. On the other hand, if the cost is paid in the perceptual domain one might expect a high level of confusion between these place contrasts.

If Steriade's hypothesis is correct one might expect better differentiation of retroflexion (vis-à-vis the other places of articulation) from the transitions alone in the VC context, both in production and perception.

We should first consider the evidence on the relative capacity of VC transitions to convey place distinctions in stops. In fact, the evidence from prior studies is mixed on this point. Sussman, Bessell, Dalston, & Majors (1997) found the locus equations (F2 value at the consonantal transition vs. F2 of the adjacent vowel) to be more variable in VC than in CV position. They review literature also showing that VC transitions afford poorer differentiation than CV transitions: Ahmed & Agrawal (1969) which, significantly, also dealt with Hindi, and Redford and Diehl (1999). The results of Wang & Bilger (1973) showed no overall difference in this regard between CV and VC contexts. On the other hand a number of earlier studies did show VC place cues more reliable perceptually than those at CV: Sharf & Hemeyer (1972), Sharf & Beiter (1974), and Ohde & Sharf (1977) found VC place cues more reliable than CV, even when played backwards, i.e., when they then appeared as CV transitions. They reject hypotheses that this result was obtained due to auditory constraints or listeners' experience, and rather advocated a theory that there was inherently more information on place in VC transitions.



Greenberg (1995) cites neurophysiological studies of neurons in the auditory cortex showing that with many neurons the response is greater for onset of a stimulus rather than its offset and speculates that this may account for the greater incidence and greater variety of syllable onsets rather than codas in languages of the world.

Studies involving the creation of medial heteroganic stop clusters through splicing, e.g., splicing the first half of [apa] and the second half of [aka] to yield the heterorganic medial cluster [apka], show that, in general, when the medial closure duration is kept short, listeners hear a single consonant, not a cluster, and the place cues are taken from the C2 consonant, i.e., the one that had the CV transitions (Dorman, Raphael, & Liberman 1979; Repp 1976, 1977a, 1977b, 1978; Fujimura, Macchi, & Streeter 1978; Ohala 1990; Beddor & Evans-Romaine 1995; Streeter & Nigro 1979). Wang (1959) demonstrated much the same by combining post-vocalic VC transitions with final stop releases from other stops. (These lab-based perceptual results mirror, in part, the sound changes exemplified by the Late Latin to Italian change of medial -C1C2- clusters to -C2C2- geminates, mentioned above.)

Reinforcing the above finding that place cues at stop release overshadow those in the VC transition are the findings of Householder

(1956) and Malécot (1958) that there are fewer errors in place judgements when final stops are released than when they are not.

In pursuing the above questions, we sought first to determine whether the VC transitions leading into the five different places of articulation in Hindi were sufficiently distinct. This is described in study 1. Then we sought to determine the perceptual distinctiveness of the place of these coda stops, both with their releases and without. This is described in study 2.

## 2. Study 1

In order to determine the character of the VC transitions leading into the 5 places of articulation in Hindi we did the following study. The specific questions addressed were: (a) Are these places cued similarly in different vocalic environments? (b) Are the five places of articulation well differentiated by formant transitions alone, i.e. without the benefit of stop or affricate releases? (c) Are the formant patterns characteristic of place similar to those found in other languages?

### *2.1 Method*

Ten male native speakers of Standard Hindi were recorded uttering syllables of the form /pVC/ where V = one of the following vowels [i ɪ ɛ u ʊ a] and C = a voiced or voiceless (unaspirated) stop that was bilabial, dental, retroflex, palatal, or velar. (Again, the palatal stop is actually a palato-alveolar affricate.) These words occasionally happened to be real words; although most were nonsense words. The test words were read in one of two different random orders in the frame [vo___aja] 'he___came'. The utterances were recorded using high quality analog portable equipment and subsequently band-pass filtered at 68 Hz to 7.8 kHz and digitized at 16.7 kHz.[2] The analysis of the recordings from three talkers was done with the aid of waveform and LPC spectral displays produced by the CSRE™ speech analysis software and related programs. The VC formant transitions (the last 100 msec of the vowel) were extracted by the CSRE™ LPC-based formant tracking algorithm. The results presented in Figure 1 are for the most part based on 9 tokens (but never less than 7) per utterance (3 tokens × 3 speakers).



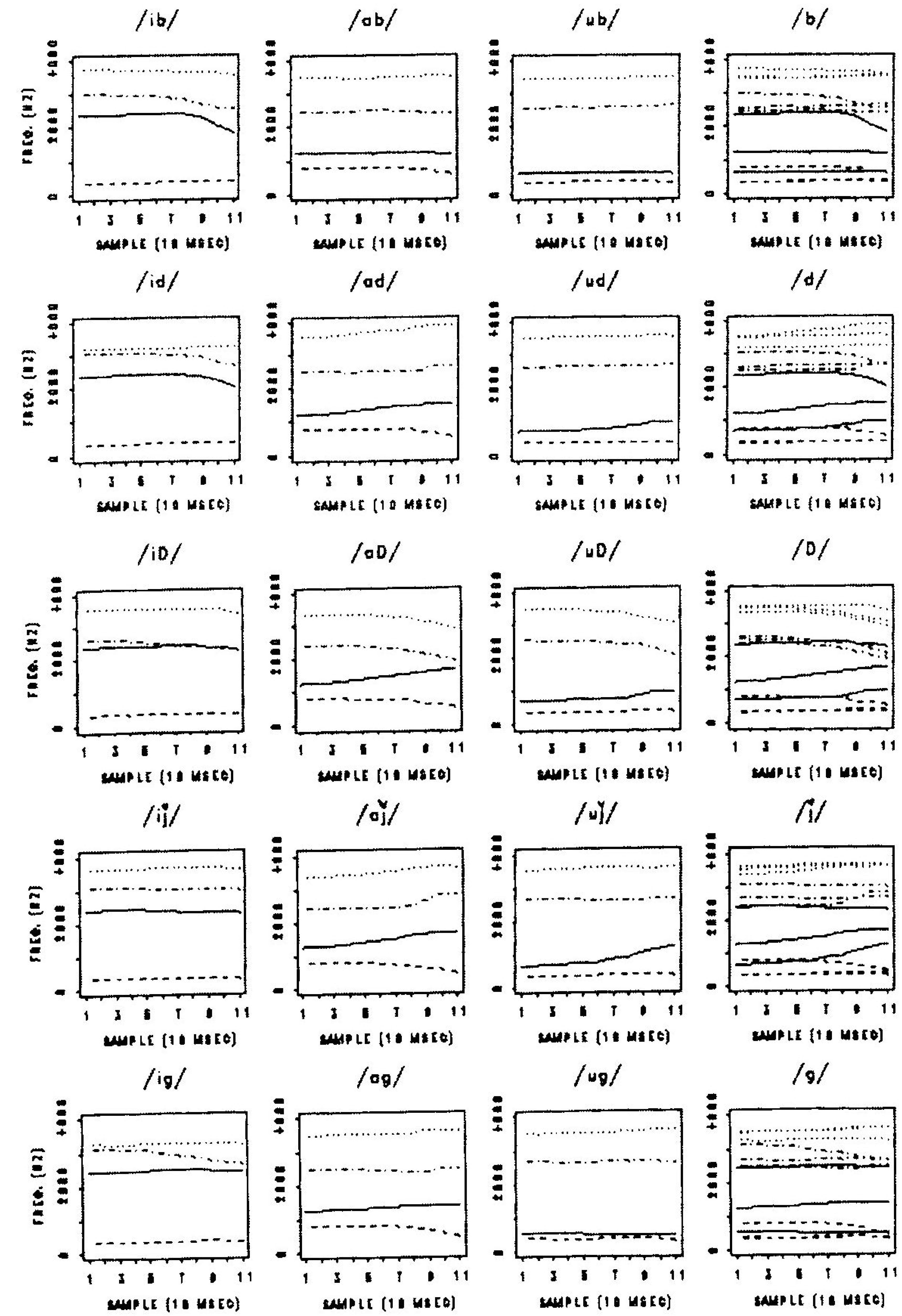


Figure 1. Formant tracks of various VC sequences in Hindi. Dashed line: F1, solid line: F2, dashdot line: F3, dotted line: F4. The symbol [D] stands for a retroflex stop. Rows: different places of articulation (labial, dental, retroflex, palatal, velar); Columns: different vowels (i, a, u); the last column superimposes the formant tracks from all three vowels. See text for furter details.

### 2.2 Results

Figure 1 gives a graphical summary of the averaged formant tracks for three vowels [i a u ] before 5 different places of voiced stop (or affricate). (The results for [ɪ ʊ ɑ] were in general similar to those of [i u a] respectively and the patterns for [ɛ ɔ] were in general interpolated between those shown here, e.g., the pattern for [ɛ] is approximately in between those for [i] and [a]. Also, although the formant tracks for voiceless consonants are not given, they were similar to those of the voiced consonants.)

Some of the formant patterns conform to expectations (based on patterns seen in other languages):

- Labials show a lowering of F2 and F3 – except after back vowels where F2 and F3 were essentially flat.
- Dentals show a transition that ends in a mid-frequency region between about 1000 to 2000 Hz.
- Palatals show an elevated F2 and F3.
- Place distinctions are most different in the environment of the open vowel [a]; they show similarities and thus potential neutralization in the environment of the high vowels [i] and [u].

Some patterns of the transitions, however, are not expected, at least based on what has been previously reported for other languages:

- Although the velar stop shows some convergence of F2 and F3 in the environment of [i], it does not show this expected pattern after [a].
- The retroflex does show convergence of F2 and F3 in all vocalic environments.



There is also a considerable amount of overlap. Thus many transitions were visually quite similar in specific vowel contexts. Specifically the following were similar:

- Labials and dentals have similar transitions after [i].
- Retroflexes and velars have similar transitions after [i].
- Dentals and palatals are similar after [a] and [u].
- Labials and velars are similar after [u].
- Dentals and velars are similar after [a].

It is also interesting to note therefore that acoustic cues for place are highly variable and context dependent – even in a language with a crowded segment inventory.

The question arises: are there other parameters which carry the necessary place cues or are these places in fact highly confusable with their differentiation requiring a stop release with its rich place cues present? A perception experiment was designed to elucidate these issues.

## 3. Study 2

In this study we sought to determine how well the place of the five stops could be perceptually differentiated both with and without the final release burst.

### *Method*

Stimuli: The recorded stimuli used was from the same data-base as the one mentioned above except one of the three speakers was different. (This was because one of the speakers from the first study often released his final stops into the vowel of the following word. Thus another speaker whose stop releases could be better isolated was selected.) In order to keep the size of the word list manageable only three vowels were used. Thus the stimuli was of the form /pVC/ where V = [i a u] and C = a voiceless unaspirated stop that was bilabial, dental, retroflex, palatal, or velar. The resulting tokens were of the sort [puk], [pap], [pitʃ], etc. (In the first study the final stops were voiced. Voiceless stops were used here because it facilitated the isolation of the final stop release. We don't believe the voicing of the final stops would have any significant effect on the overall trends in the results.) The data, digitized at 16 kHz after low-pass filtering at 8 kHz, were further processed to yield two versions of each word type: the "whole word" version, and a gated version that excluded the signal after the halfway point in the silence of the final stop, i.e., excluding the final burst. No part of the frame sentence was included. Thus the stimulus list contained 90 items: 3 vowels × 5 stops × 2 conditions × 3 subjects = 90. This 90 item word list was randomized and recorded on a tape with an interstimulus of 4 s. After every 5 tokens a 10 second gap was given (to allow subjects to find their way in case they lost their place on the answer sheet – although none did). Preceding the 90 item word list was a 10 item practice session to familiarize the subjects with the way the stimuli would sound, how long they would have to respond, and with the way they were to mark their answer sheet which gave the candidate answers in the Devanagari script. These 10 tokens were similar to the words for the main test except they were prepared from /pVC/ tokens where the C was voiced. Five were in the "whole word" format and 5 in the gated format. For these, too, there was a 4 s. gap between tokens. Between the 10 tokens of the practice session and the tokens of the main test there was a 20 second gap. This test was administered over headphones via a high-quality portable tape playback system in 21 subjects' homes in India.

Subjects and Procedures: The subjects were between the ages of 13-60; they could all read the Devanagari orthography. They were paid for their participation. They were provided with an answer sheet which listed the five final consonants [k tʃ ʈ t̪ p] in the Devanagari orthography for each of the 90 items of the test session as well as the practice session. The instructions to the subjects given orally in Hindi were as follows: The purpose of this experiment is to examine the intelligibility of speech in noisy conditions. You will be hearing words some of which are real words and some nonsense words. Some of the words will sound quite clear, for others the end might not be so clear. You should listen carefully and circle on the answer sheet the consonant that you think the word ends with. You **must** circle one consonant (and only one consonant) for each item even if you have to guess at



what it is. There will be 4 second gaps between each item and longer gaps after every 5 items in order to help you find your place. We will start with a practice session in order for you to get used to the format. After the practice session if there were no questions we proceeded with the main test.

### *3.2. Results*

| **i-gated** | **p** | **t̪** | **ʈ** | **tʃ** | **k** | **i-whole** | **p** | **t̪** | **ʈ** | **tʃ** | **k** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **p** | **90.5** | 7.9 | 0.0 | 0.0 | 1.6 | | **95.2** | 4.8 | 0.0 | 0.0 | 0.0 |
| **t̪** | 54.0 | **36.5** | 6.3 | 0.0 | 1.6 | | 4.8 | **71.4** | 23.8 | 0.0 | 0.0 |
| **ʈ** | 19.0 | 39.7 | **36.5** | 1.6 | 3.2 | | 1.6 | 0.0 | **92.1** | 4.8 | 1.6 |
| **tʃ** | 20.6 | 39.7 | 11.1 | **11.1** | 17.5 | | 0.0 | 1.6 | 17.5 | **76.2** | 4.8 |
| **k** | 3.2 | 11.1 | 0.0 | 0.0 | **81.0** | | 0.0 | 0.0 | 0.0 | 0.0 | **95.2** |

Table 3. Confusion matrix (in percent) for /iC/: left: with release gated; right: with release intact. Correct response along diagonal.

| **u-gated** | **p** | **t̪** | **ʈ** | **tʃ** | **k** | **u-whole** | **p** | **t̪** | **ʈ** | **tʃ** | **k** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **p** | **92.1** | 6.3 | 0.0 | 0.0 | 1.6 | | **84.1** | 6.3 | 9.5 | 0.0 | 0.0 |
| **t̪** | 3.2 | **90.5** | 4.8 | 1.6 | 0.0 | | 31.7 | **65.1** | 3.2 | 0.0 | 0.0 |
| **ʈ** | 9.5 | 17.5 | **66.7** | 4.8 | 1.6 | | 0.0 | 0.0 | **95.2** | 4.8 | 0.0 |
| **tʃ** | 3.2 | 50.8 | 3.2 | **38.1** | 4.8 | | 0.0 | 0.0 | 0.0 | **95.2** | 4.8 |
| **k** | 49.2 | 0.0 | 3.2 | 1.6 | **42.9** | | 1.6 | 0.0 | 1.6 | 0.0 | **92.1** |

Table 4. Confusion matrix (in percent) for /uC/: left: with release gated; right: with release intact. Correct response along diagonal.

| **a-gated** | **p** | **t̪** | **ʈ** | **tʃ** | **k** | **a-whole** | **p** | **t̪** | **ʈ** | **tʃ** | **k** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **p** | **92.1** | 4.8 | 0.0 | 1.6 | 1.6 | | **93.7** | 4.8 | 1.6 | 0.0 | 0.0 |
| **t̪** | 9.5 | **81.0** | 6.3 | 0.0 | 1.6 | | 0.0 | **90.5** | 9.5 | 0.0 | 0.0 |
| **ʈ** | 1.6 | 6.3 | **87.3** | 4.8 | 0.0 | | 0.0 | 0.0 | **95.2** | 4.8 | 0.0 |
| **tʃ** | 3.2 | 54.0 | 7.9 | **25.4** | 9.5 | | 0.0 | 0.0 | 0.0 | **95.2** | 4.8 |
| **k** | 3.2 | 19.0 | 0.0 | 0.0 | **73.0** | | 0.0 | 0.0 | 0.0 | 0.0 | **95.2** |

Table 5. Confusion matrix (in percent) for /aC/: left: with release gated; right: with release intact. Correct response along diagonal.

ter /a/ and /u/, labials and velars after /u/, and velars and dentals after /a/. An expected confusion between retroflexes and velars after /i/ did not manifest itself – possibly because the retroflexes, unlike the velars, tended to induce a characteristically low F3 throughout the preceding vowel.

There was also some confusions that were not predicted, at least based on a visual examination of the formant transitions, namely, palatals and retroflexes were confused with both dentals and labials after /i/ and retroflexes were confused with dentals after /u/.

Over all three vocalic environments, the rank order of percent correct identification of place was, from more to less: labial (91.6%), dental (69.3%), velar (65.6%), retroflex (63.5%), and palatal (24.9%).

The degree to which place perception was improved, on average, by the addition of the final stop release varied considerably:

| | |
|---|---|
| palatals | 64% |
| retroflex | 31% |
| velar | 29% |
| dental | 6.3% |
| labial | – .57% |

This ranking is roughly in the same order as the typical prominence and robustness of the releases of these stops. The palatal (affricate) is known to have an intense characteristic fricative release and the back-articulated velar and the retroflex stops would also tend to have relatively intense bursts by virtue of their having downstream resonators to amplify their sound.

Regarding Steriade's (1997) claim: it is true that the retroflex has robust VC transitions after [a], and corresponding to this the percent correct identification in the gated conditions was a high 87.3%. However, after [i] this is not true: the transitions are less robust (Dave 1977; M. Ohala 1995) and the rate of correct identification was only 36.5%. Thus claims about the robustness of the retroflex place distinction in VC context need to take into consideration the nature of the V. Additionally, another factor may have to be taken into account in evaluating the relative salience of retroflexes' cues in some VC positions. It is possible that via sound change the presence of a syllable final retroflex consonant has yielded a different preceding vowel, e.g., in Dave's (1977) study the /ə/ before a retroflex may no longer be the same as the /ə/ before other consonant places. The data presented by Dave (1977) for Gujarati retroflexes suggests, for example, that the preceding vowel itself, not just the transitions, may differ from the same vowel before dentals. Similarly, in the history of English a post-vocalic /r/ has had major influences on preceding vowels (Dobson 1968). One example is that after metathesis changed *brid* to 'bird' the vowel underwent lowering and centralization. (We cite this as an example of how a consonantal context can influence the phonological quality of vowels. It is not necessarily directly relevant to retroflexes' influence since the original /r/ in English may not have been retroflex.) If such a vowel change is involved, the identification of final retroflexes may in some cases be helped by a correlated difference in vowel quality, not just by the consonantal transitions per se.



Additionally, the results obtained after [ i ] would not have been predicted on the basis of examining the formant transitions. Anderson's (1995) results that post-alveolars were identified correctly more often in intervocalic position than word-initial may not necessarily show that these sounds' VC transitions are better cues to that place than the CV transitions; it could simply mean that the combination of VC and CV cues (since both are present intervocalically) are better than CV cues by themselves.

One final note of caution: although it seems reasonable to think that VC intervocalically (the position reported on by Anderson (1997), Dave (1977) and Steriade (1997) and VC in final position (used in the present study) would behave similarly for formant transitions and how they are perceived, this might not be so and still needs to be examined.

**Acknowledgments**
This is a revised version of two previous papers. The first study was previously reported as "Acoustic study of VC transitions for Hindi stops" *Proceedings of the XIIIth International Congress of Phonetic Sciences, Stockholm, August 13-19, 1995* by the first author. For this study she thanks Anvita Abbi and V.V. Sahni for their help in the use of the facilities at Jawahar Lal Nehru University and T. Nearey and T. Wells for help in using the speech analysis system at the University of Alberta. That research was supported by a sabbatical leave from San José State University and a research grant from the University of Alberta. The second study appeared earlier as a report by both of the present co-authors as "Correlation between consonantal VC transitions and degree of perceptual confusion of place contrast in Hindi" *Proceedings, International Conference on Spoken Language Processing, Sydney, 30 November-4 December 1998*. For that research we thank G. Ying for his help. Finally, we thank our subjects in India for their time and attention.

## References

Ahmed, R. & S.S. Agrawal. 1969. Significant features in the perception of (Hindi) consonants. J. Acoust. Soc. Am. 45.758-63.

Anderson, V. 1997. The perception of coronals in Western Arrernte. Proc. Eurospeech, Rhodes, Greece, 22-25 September 1997 1.389-92.

Beddor, P.S. & D. Evans-Romaine. 1995. Acoustic-perceptual factors in phonological assimilations: A study of syllable-final nasals. Rivista di Linguistica 7.145-74.

Dave, R. 1977. Retroflex and dental consonants in Gujarati. A palatographic and acoustic study. Ann. Rep. Inst. Phonetics, Univ. Copenhagen 11.27-156.

Dobson, E.J. 1968. English pronunciation 1500-1700. 2nd ed. Oxford: Oxford University Press.

Dorman, M., L.J. Raphael & A.M. Liberman. 1979. Some experiments on the sound of silence in phonetic perception. J. Acoust. Soc. Am. 65.1518-32.

Fischer-Jorgensen, E. 1967 [1954]. Acoustic analysis of stop consonants. Readings in acoustic phonetics, ed. by Ilse Lehiste, 137-54. Cambridge, MA: MIT Press.



Fujimura, O., M.J. Macchi & L.A. Streeter. 1978. Perception of stop consonants with conflicting transitional cues: a cross-linguistic study. Language & Speech 21.337-46.

Ghazeli, S. 1977. Back consonants and backing coarticulation in Arabic. Unpub. Doc. Diss., University of Texas, Austin.

Greenberg, S. 1995. The ears have it: The auditory basis of speech perception. Proc. 13th Int. Congr. Phonetic Sciences, ed. by K. Elenius & P. Branderud, 31-41.

Householder, F. 1956. Unreleased ptk in American English. For Roman Jakobson, ed. by M. Halle, H.G. Lunt, H. McLean & C.H. van Schooneveld, 235-44. The Hague: Mouton.

Kawasaki, H. 1982. An acoustical basis for universal constraints on sound sequences. Doc. Diss., Univ. Calif., Berkeley.

Lindblom, B. 1984. Can the models of evolutionary biology be applied to phonetic problems? Proc. 10th Int. Congr. Phonetic Sciences, ed. by M.P.R. van den Broecke & A. Cohen, 67-81. Dordrecht: Foris.

Lindblom, B. 1989. Phonetic invariance and the adaptive nature of speech. Working models of human perception, ed. by B.A.G. Elsendoorn & H. Bouma, 139-52. London: Academic Press.

Lindblom, B. & I. Maddieson. 1988. Phonetic universals in consonant inventories. Language, speech, and mind: Studies in honor of Victoria A. Fromkin, ed. by L.M. Hyman & C.N. Li, 62-80. London: Routledge.

Maddieson, I. 1984. Patterns of sounds. Cambridge: Cambridge University Press.

Malécot, A. 1958. The role of releases in the identification of released final stops. Language 34.370-80.

Ohala, J.J. 1975. Phonetic explanations for nasal sound patterns. Nasálfest: Papers from a symposium on nasals and nasalization, ed. by C.A. Ferguson, L.M. Hyman, & J.J. Ohala, 289-316. Stanford: Language Universals Project.

Ohala, J.J. 1985. Around flat. Phonetic linguistics. Essays in honor of Peter Ladefoged, ed. by V. Fromkin, 223-41 Orlando FL: Academic Press.

Ohala, J.J. 1990. The phonetics and phonology of aspects of assimilation. Papers in Laboratory Phonology I: Between the grammar and the physics of speech, ed. by J. Kingston & M. Beckman, 258-75. Cambridge: Cambridge University Press.

## Notes

1 We use the term 'space' as a convenient metaphor but do not endorse many of the concepts associated with it (J. Ohala 1997).

2 The recordings were made in New Delhi, India, in the Language Laboratory of the Jawahar Lal Nehru University.

3 The rows occasionally add up to more or less than 100% due to round-off error.