

“Perceiving the Advantage”:

**the effect of prior L2 exposure versus perception training
on L1 English speakers’ acquisition of Hindi aspirated consonants**

Anupama Bhatt

Advisor: Dr. Keith Johnson

Acknowledgments

I would like to sincerely thank my advisor, Dr. Keith Johnson, for his help, advice, and patience from the beginning to the end of this project. Thank you for your reassurance and encouragement!

I would also like to thank my second reader, Dr. Line Mikkelsen, for taking the time to read my drafts and give helpful suggestions.

Finally, to you, Reader:

If you are:

- a) a friend or a family member, I thank you for reading this (possibly against your will).
- b) an innocent, curious stranger, I hope the next forty pages do not put you to sleep. Make sure to enjoy this paper in a dry, academic way.

1. Introduction

An important component of second-language (L2) acquisition is the ability to accurately perceive the language's unfamiliar sounds. A native (L1) English speaker might have trouble producing a “foreign” sound if he/she has never heard that sound before, while a native English speaker with prior exposure to the “foreign” sound would recognize it, and therefore be able to produce it more easily. The advantage of prior L2 exposure has been studied with childhood ‘overhearers’ who, once re-introduced to the language in adulthood, could produce fairly accurate vowel and consonant articulations when compared to L1 speakers of the language and late L2 learners (Au, Knightly, Jun, & Oh, 2002).

Thus, it has been established that prior L2 exposure provides a certain perceptual advantage in adulthood. Other studies have been conducted in which subjects unfamiliar with a L2 are trained to perceive sounds in that language. Training can even employ different teaching methods—auditory exposure versus verbal instruction, for example—to see which method helps the subject more accurately perceive the unfamiliar sounds (Catford & Pisoni, 1970).

Therefore, a new question arises: if a subject with no prior L2 exposure to a language is trained in perception and is compared to someone with prior L2 exposure, would the latter then still retain an advantage? My hypothesis is yes: prior L2 exposure still provides the greatest advantage to a speaker. However, training can still significantly improve a subject's perception of “foreign” sounds.

I am conducting this study to test this hypothesis, using Hindi as the language of focus.

Hindi stems from the Indo-Iranian branch of the Indo-European language family. It is largely derived from Sanskrit, and is written using the Devanagari script. Hindi has a class of consonants known as breathy consonants. These are, essentially, consonants with aspiration. One essential disparity between Hindi and English phonology is the phonemic distinction in Hindi between plain and aspirated consonants. Furthermore, Hindi contains not only voiceless aspirated [h] consonants but voiced aspirated [ɦ] consonants. Both voiceless and voiced aspirated consonants can appear in word-initial, word-medial, and word-final position.

In this study, L1 English speakers serve as the source of data to test my hypothesis concerning prior L2 exposure versus perception training. All subjects have prior L2 exposure to a language, but only some have prior L2 exposure to Hindi. I am using the Auditory and Articulatory training methods employed by Catford and Pisoni (1970) to teach perception of Hindi sounds to the non-Hindi learners. Whereas Catford and Pisoni tested and trained subjects with “exotic” sounds from languages like Russian, German, and French, I am using voiceless and voiced plain and aspirated Hindi consonants as stimuli. Catford and Pisoni used sounds such as the close back unrounded tense vowel [u], the voiceless dorso-uvular stop [q], and the close front rounded tense vowel [y]. I am using sounds that include the voiceless aspirated velar stop [k^h], the voiced aspirated alveopalatal affricate [dʒ^h], and the voiced aspirated bilabial stop [b^h].

I have expanded Catford and Pisoni’s subject population by incorporating Hindi learners into the study. Due to their prior L2 exposure to the language of focus, the Hindi learners act as the control group and are compared to the non-Hindi learners, who are “taught” perception using the Articulatory and Auditory training methods. Results from

the two training groups are expected to replicate the results of Catford and Pisoni's experiment: the Articulatory training group is expected to surpass the Auditory training group in both discrimination and production. However, I expect that—overall—the Hindi learners will obtain the highest results in both tasks. This will support the hypothesis that a group with prior L2 exposure to a language of focus has an advantage over a group without prior L2 exposure to that language.

In the next section, I will discuss previous studies that have influenced my experiment. Section 3 is an explanation of my method, including my subject population and preparation of testing materials. Section 4 details the techniques I employed in the training sessions. Section 5 summarizes when and where testing and training took place. Results are discussed in Section 6, followed by a conclusion. The Appendices contain the Devanagari alphabet, the L2 background of each subject, and data charts of individual and average percent-accuracy on the discrimination and production tasks.

2. Previous Studies

The intriguing phenomenon of L2 acquisition has been analyzed and studied in various ways. One attempt has been to “teach” perception by training individuals in “exotic” sounds (Catford & Pisoni, 1970). Fourteen subjects from the University of Michigan were taught using the Articulatory and Auditory training methods. The Articulatory method involved systematic instruction in *how* to pronounce certain sounds, which the Auditory method involved repeated auditory exposure to the sounds with frequent comparisons between foreign and familiar sounds. (Familiar sounds were the consonants [s, ʃ, l, k] and the cardinal vowels [a, e, i, o, u].) Articulatory subjects only received auditory exposure through their own mimicry of the sounds. They arrived at

correct pronunciation through their own articulation, with minimal feedback from the instructor.

As for the Auditory subjects, no pronunciation instruction was provided. Subjects mimicked each new sound, and were given minimal feedback.

After training, subjects in both groups were given an auditory identification test with phonetic symbols that they had previously learned, followed by a production exam in which they were scored by two phoneticians on a scale of 0-2 (0-nowhere close; 1-fairly good; 2-perfect or near-perfect).

Results showed that the Articulatory training method significantly improved subject perception over the Auditory training method. The critical discovery linked to this study was that a lack of instruction in sound production—as experienced by the auditory group—is damaging to a subject’s perception of a foreign language. A subject will obtain better results on a perception test if he/she knows *how to produce the sound*—as is taught in the Articulatory method.

Another experiment explored the theory that childhood overhearers of a L2 retain a perceptual advantage in adulthood over late L2 learners (Au, Knightly, Jun, & Oh, 2002). Eleven overhearers—hearers of informal Spanish for several hours a week for at least three years between birth and age 6—and many late L2 learners—people who had minimal exposure to Spanish until regular language instruction at age 14—were compared to L1 Spanish speakers. Accent and vocabulary were tested using thirty-six Spanish sentences and forty childhood slang terms as stimuli. Childhood overhearers produced more accurate vowel and consonant articulations than late L2 learners, and their accent resembled that of the L1 Spanish speaker. Thus, the study showed that childhood

overhearing does benefit adults when re-exposed to the language. Au, et al. based their theory on the suggestion that early exposure to a L2 establishes a long-lasting mental representation of that language (Chomsky, 1981). It is this “mental representation” that is accessed when a childhood overhearer is re-exposed to the L2 in adulthood.

An experiment very similar to Catford and Pisoni’s 1970 study examined native Japanese speakers’ acquisition of the English /r-l/ minimal pair using an auditory training method (Akahane-Yamada, Tohkura, Bradlow, & Pisoni, 2000). The twenty-three subjects who participated were between the ages of 18 and 22 and had studied English since junior high school, though none had lived abroad or had received special training in English conversation. They were divided randomly into two groups. One group served as the control, without training. The other group was trained between the pre- and post-test (based on the Strange & Dittmann, 1984 pretest/post-test design) with a two alternative forced choice (2AFC) task. The same minimal-pair stimuli from the pre- and post-tests were used in the training session, with the addition of auditory feedback (a chime for “correct”, a buzzer for “incorrect”). Subjects were also recorded while reading a set of fifty-five English /r-/l/ minimal pairs. Ten American English speakers scored the recordings.

The subjects who had undergone training showed improvement over the control group in both perception and production, even after gaps of three months and six months from the date of preliminary testing. The trained subjects’ production results from the post-test to the three-month and six-month gap remained superior to their pre-test production results. The control group, however, did not show significant improvement

between the pre-test and post-test, nor after a three-month gap. They were not evaluated for the six-month follow-up.

Davis and Kim (2001) explored the effects of various techniques used in training methods. The focus of the experiment was on the benefit of visible speech and its correlation to subjects' foreign language perception. The study built on the theory that subjects familiar with a foreign language will fare better on a discrimination task than subjects who are unfamiliar with that language (Werker & Logan, 1985). Linguists have compared the data from Werker and Logan to data from studies in which subjects unfamiliar with a language were trained in perception of its sounds. Davis and Kim took it further to ask whether visible speech during training—that is, letting the instructor's mouth be seen while he/she produces sounds—can speed up the acquisition process for subjects unfamiliar with the language.

Results showed that yes, subjects unfamiliar with Korean benefited from seeing the instructor's mouth. A task was used that was similar to “delayed shadowing” (Balota & Chumbley, 1985), where the subject was allowed to mimic the sound only after the third repetition, to test for sound retention. This was chosen over “immediate shadowing”—mimicking after the first presentation of the stimulus (Reisberg, McLean, & Goldfield, 1987)—because “immediate shadowing” was thought to demonstrate use of short-term memory. Goldinger (1998) referred to this phenomenon as the “echo content” of memory. Since Davis and Kim were focusing on more permanent retention of the stimuli, they chose to have the subjects mimic the stimulus only after the third repetition, thereby creating a task similar to “delayed shadowing”.

I incorporated “delayed shadowing” into the auditory discrimination tasks and into both the Articulatory and Auditory training sessions of the study reported below. The auditory discrimination tasks were based on the pre-test/post-test design (Strange and Dittmann, 1984) and on a same-different paradigm. I used Catford and Pisoni’s Articulatory and Auditory training methods as a template off of which I modeled the training sessions. The subjects involved in the study report below were not childhood overhearers, but university students who had studied various foreign languages since high school or earlier. Only a select few were Hindi learners.

3. Method

Participants

Fifteen undergraduate and graduate students participated in the experiment. Three were graduate students, and twelve were undergraduates. All were from the University of California, Berkeley (UC Berkeley). The subjects were divided into two groups based on their language background: Group A: prior exposure to foreign language, in college or earlier, but no prior exposure to Hindi. Group B: at least one semester of Hindi at UC Berkeley. Group A had ten participants—two males and eight females; Group B had five—three males and two females. Gender was not a focus of the experiment, nor did I notice any significant perceptual advantage of male subjects over female subjects, or vice versa.

Group B was not required to undergo any training, as they were the control group for the experiment. They were paid \$10 total for participation. Group A was compensated \$15 total for their participation, because they were required to attend a training session. Subjects were paid \$5 after each session, or the total owed to them at the end of the study, based on their preference.

Materials

I incorporated the ten main voiceless and voiced aspirated consonants in modern Hindi into all the tasks and training sessions. The consonants are listed below, in the order in which they appear in the Devanagari alphabet.

(Please refer to Appendices, Table 1 to see the full Devanagari alphabet.)

[k^h] – voiceless aspirated velar stop

[g^h] – voiced aspirated velar stop

[tʃ^h] – voiceless aspirated alveopalatal affricate

[dʒ^h] – voiced aspirated alveopalatal affricate

[tʰ] – voiceless aspirated retroflex stop

[dʱ] – voiced aspirated retroflex stop

[tʰ] – voiceless aspirated dental stop

[dʱ] – voiced aspirated dental stop

[p^h] – voiceless aspirated bilabial stop

[b^h] – voiced aspirated bilabial stop

Before making sound recordings, I created a list of thirty actual Hindi words. Each word incorporated one of the ten consonants in word-initial, word-medial, and word-final position.

There were some exceptions: [dʱ] and [p^h] become [ɽ^h] and [f] in word-final position, therefore I did not include a word for them in the “word-final” column.

/ dʱ/ → [ɽ^h] / __#

/ p^h/ → [f] / __#

The words are listed below:

	word-initial	word-medial	word-final
[k ^h]	[k ^h ir](n.fsg.) ‘rice pudding’	[sik ^h na] (v.) ‘to learn’	[ā:k ^h] (n.fsg.) ‘eye’
[g ^h]	[g ^h ər] (n.msg.) ‘house’	[pɪg ^h əlna](v.) ‘to melt’	[ba:g ^h] (n.msg) ‘tiger’
[ʃ ^h]	[ʃ ^h a:ja] (n.fsg.) ‘shade’	[pɪʃ ^h lɛ] (adj.) ‘previous’	[mũ:ʃ ^h] (n.fsg.) ‘mustache’
[ɖʒ ^h]	[ɖʒ ^h ɛlna] (v.) ‘to endure’	[bɪɖʒ ^h ra] (n.msg.) ‘grain’	[sā:ɖʒ ^h] (n.fsg.) ‘evening’
[t ^h]	[t ^h i:k] (adj.) ‘straight, correct’	[pəʃ ^h a:r] (n.msg.) ‘plateau’	[a:t ^h] (adj.) ‘eight’
[d ^h]	[d ^h ai] (adj.) ‘two and a half’	[mɛnd ^h ək] (n.msg.) ‘frog’	
[ʈ ^h]	[ʈ ^h oɾa] (adj.) ‘small amount’	[sā:ʈ ^h əl] (n.fsg.) ‘thigh’	[sa:ʈ ^h] (prep.) ‘with’
[ɖ ^h]	[ɖ ^h i:ɾɛ] (adv.) ‘slowly’	[ʌɳɖ ^h ɛra] (n.msg.) ‘darkness’	[sā:ɖ ^h] (n.fsg.) ‘crevice’
[p ^h]	[p ^h ɪr] (adv.) ‘then’	[gəp ^h a:ɛ] (n.fpl.) ‘caves’	
[b ^h]	[b ^h ɛ:ɖʒna] (v.) ‘to send’	[ləgb ^h əg] (adv.) ‘approximately’	[ʃʊb ^h] (adj.) ‘auspicious’

I had P.G., a native Hindi speaker, read these words using a hand-held microphone in the UC Berkeley Phonology Laboratory sound booth (52A Dwinelle Hall). The sounds were digitized to computer files using Wavesurfer v. 1.8.5 (Sjölander and Beskow, 2005). I also had P.G. record consonants in isolation. By “isolation”, I mean that the consonant is recorded as it is pronounced in the Devanagari alphabet—that is, with the inherent vowel [ə]. When Hindi speakers recite the Devanagari alphabet, they say [kə, k^hə, gə, g^hə], just as English speakers say [ɛɪ, bi, si, di]. I, myself, recorded a supplementary set of sounds that contrasted plain consonants (without aspiration) to their

aspirated counterparts ([k] and [k^h], for example). I used Wavesurfer v. 1.8.5 for these recordings as well.

Next, I created a preliminary discrimination task (a pre-test) with a same-different paradigm to establish subjects' initial levels of Hindi consonant perception. In the test, subjects heard pairs of Hindi consonants in isolation. These sounds were selected from P.G.'s and my recordings and played to subjects on a computer in the Phonology Laboratory. The task was to distinguish aspirated consonants—both voiceless and voiced—from plain consonants. After the third repetition of the same pair, subjects had to mark on a sheet of paper whether they thought each pair was “same” or “different”. (This demonstrates my incorporation of “delayed shadowing”.)

There were three pairs of sounds for each of the ten aspirated consonants, making thirty pairs total. The pairs consisted of a) the plain consonant repeated twice—[kə] and [kə]; b) the aspirated consonant repeated twice—[k^hə] and [k^hə]; or c) one plain consonant and its aspirated counterpart—[kə] and [k^hə]. Two-thirds of the pairs were “same” and one-third was “different”. I randomized the order of these pairs.

(Note: [k^hə] is used as an example; the same rule applied to the other nine voiceless and voiced aspirated consonants.)

The next session of the study was training, which is explained in detail in Training Techniques, directly following this section.

The last session of the experiment involved a post-test discrimination task and a production exam, to evaluate subjects' improvement in perception and production of Hindi breathy consonants.

The post-test was the same format as the pre-test.

The production exam was a revised version of the exam in Catford and Pisoni's study. Instead of various exotic sounds such as a voiceless dorso-uvular stop [q] or close front rounded tense vowel [y], I tested subjects on their pronunciation of the ten main Hindi aspirated consonants (listed above in phonetic translation). The production exam was incorporated into the experiment to provide an 'articulatory' evaluation of subjects' facility with Hindi breathy consonants, since the discrimination tasks were purely auditory. However, it also involved an element of auditory perception, since subjects were required to listen and mimic my production of each consonant.

I created thirty made-up words, designed to test subjects' production of each consonant in word-initial, word-medial, and word-final position. I did not want subjects to get confused by the presence of other consonants in the word, so I chose not to use the real Hindi words from the chart on page 11. Rather, I made up words that contained each breathy consonant in CV, VCV, and VC positions. To add variety, and an added level of difficulty for the Hindi learners, I reversed the order of the ten consonants. As shown in the chart below, I started with bilabial consonants rather than velar consonants, and worked my way backwards (in terms of place of articulation) to velar stops. I created the CV sets with the high front tense unrounded vowel [i], the VCV sets had the low back tense unrounded vowel [a], and the VC sets had the high back tense rounded vowel [u].

	word-initial	word-medial	word-final
[p ^h]	[p ^h i]	[a:p ^h a:]	[u:p ^h]
[b ^h]	[b ^h i]	[a:b ^h a:]	[u:b ^h]
[t ^h]	[t ^h i]	[a:t ^h a:]	[u:t ^h]
[d ^h]	[d ^h i]	[a:d ^h a:]	[u:d ^h]
[t ^h]	[t ^h i]	[a:t ^h a:]	[u:t ^h]
[d ^h]	[d ^h i]	[a:d ^h a:]	[u:d ^h]
[ʃ ^h]	[ʃ ^h i]	[a:ʃ ^h a:]	[u:ʃ ^h]
[dʒ ^h]	[dʒ ^h i]	[a:dʒ ^h a:]	[u:dʒ ^h]
[k ^h]	[k ^h i]	[a:k ^h a:]	[u:k ^h]
[g ^h]	[g ^h i]	[a:g ^h a:]	[u:g ^h]

4. Training techniques

In their experiment, Catford and Pisoni compared the effect of the Articulatory method and Auditory method of training on subjects' perception of "exotic" sounds. In Articulatory training, subjects were verbally instructed by Catford in how to produce the sounds (*"keep your glottis closed, and while the glottis is closed, [try] to produce a sort of k-sound"*). Catford made minimal productions of the sounds himself. Auditory training subjects listened repeatedly to the "exotic" sounds and then mimicked them. Both training sessions included some written script.

In my study, Group A subjects—those with prior L2 exposure but no prior exposure to Hindi—were divided at random into two training groups. Four subjects were assigned to the Articulatory training group and the remaining six made up the Auditory

training group. The reason for the unequal distribution was that I wanted to challenge the Articulatory training method. If it was indeed superior to the Auditory method—it was, according to the result of Catford and Pisoni’s experiment—then Articulatory trainees should still be able to obtain superior results, even with a smaller population than the Auditory trainees.

I prepared a lesson plan that fleshed out Catford and Pisoni’s techniques with a step-by-step progression from plain voiceless stops to voiced aspiration. Hindi consonants were compared and contrasted with the following familiar consonants: [p, t, k, b, d, g]. The voiceless stops [p, t, k] were used to illustrate word-initial voiceless aspiration and removal of aspiration through word comparisons such as *pill* vs. *spill*, *tan* vs. *Stan*, and *cool* vs. *school*. The voiced stops [b, d, g] were used to help subjects understand voiced aspiration in Hindi by combinations with the word *hill*, such as *b...hill* to *b-hill* to [b^hɪl]. Both training sessions covered the following points:

1) Voiceless versus voiced stops in English:

[p] [t] [k]

[b] [d] [g]

Subjects learn to articulate English stops as follows:

[pə...pə...pə...] [bə...bə...bə...]

[tə...tə...tə...] [də...də...də...]

[kə...kə...kə...] [gə...gə...gə...]

(voiceless) (voiced)

2) Aspiration in English and where it occurs:

pill vs. *spill*, *tan* vs. *Stan*, and *cool* vs. *school*

Subjects review words in English that aspirate voiceless consonants word-initially, but not when preceded by another consonant, like [s].

3) How to produce plain voiceless stops:

s...pill; s...tan; s...chool

This method separates the word-initial consonant [s] from the voiceless consonants [p, t, k]. English speakers pronounce [p, t, k] without aspiration when they are preceded by another consonant, like [s]. Therefore, eliminating [s] before [p, t, k] can help subjects to produce the voiceless stops as plain consonants in word-initial position, rather than as aspirated consonants. By separating [s] from the rest of the word, subjects understand that plain voiceless stops can occur word-initially, as they do in Hindi. Below are three examples:

[pɪla] (adj.) ‘yellow’

[ʈalna] (v.) ‘to delay, put off’

[kəmai] (n.fsg.) ‘earnings’

4) Emphasizing voiceless aspiration

Subjects return to word-initial voiceless aspiration, and are taught that plain voiceless stops and aspirated voiceless stops are distinct phonemes in Hindi. Both plain and aspirated voiceless stops can also appear in word-medial and word-final positions.

The following four words contain an aspirated voiceless stop in either word-medial or word-final position:

[sikʰna] (v.) ‘to learn’

[pɪʃ^hlɛ] (adj.) ‘previous’

[gʊp^ha:ɛ] (n.fpl) ‘caves’

[ã:k^h] (n.fsg.) ‘eye’

5) Dental and retroflex stops

[t̪] [d̪] [t̠] [d̠]

Subjects are introduced to voiceless and voiced dental and retroflex stops, which are distinct phonemes in Hindi and can occur in word-initial, word-medial, and word-final position.

6) Voiceless aspiration → voiced aspiration

b...hill → *b..hill* → *b—hill* → [b^hɪl]

ɖ...hill → *ɖ..hill* → *ɖ—hill* → [ɖ^hɪl]

ɖ̠...hill → *ɖ̠..hill* → *ɖ̠—hill* → [ɖ̠^hɪl]

g...hill → *g..hill* → *g—hill* → [g^hɪl]

Subjects learn that voiced aspiration also exists word-initially in Hindi. Using *hill*, they learn to aspirate voiced consonants, including voiced dental and retroflex stops. Though voiced aspirated stops also occur word-medially and word-finally in Hindi, I did not include examples in the training session, for the sake of both simplicity and time efficiency.

This lesson plan was incorporated into the Auditory training session through sound files that were spliced to create these sequences. They were played to the subjects over loudspeakers at a computer in the Phonology Laboratory. *pill* vs. *spill*, for example, was played multiple times to subjects to emphasize addition and removal of aspiration to

the voiceless bilabial plosive [p]. The same method was used with *tan* vs. *Stan* and *cool* vs. *school*.

[pʰɪl.spɪl.pʰɪl.spɪl.pʰɪl.spɪl]

[tʰɛn.stɛn.tʰɛn.stɛn.tʰɛn.stɛn]

[kʰul.skul.kʰul.skul.kʰul.skul]

Following each sequence of words—[pʰɪl.spɪl.]—I had the subjects mimic the sounds, both individually and as a group. Here again is an incorporation of “delayed shadowing”: the subjects mimicked the sounds only after the third repetition, as demonstrated in phonetic translation above. I then played sound files where [s] had been elongated: [s:.pɪl], [s:.tɛn] and [s:.kul]. I had them try to mimic the sounds. They tried to produce *pill*, *tan*, and *cool* without aspiration: [pɪl], [tɛn], [kul]. I used sound files of the four Hindi words from part 4 of the lesson plan above to demonstrate the production of voiceless aspirated Hindi consonants in word-medial, and word-final positions. Next, I taught voiced aspiration using a sound file of *hill* that had been recorded and spliced with [b, ɖ, d, g]. All splicing was done using Wavesurfer v. 1.8.5.

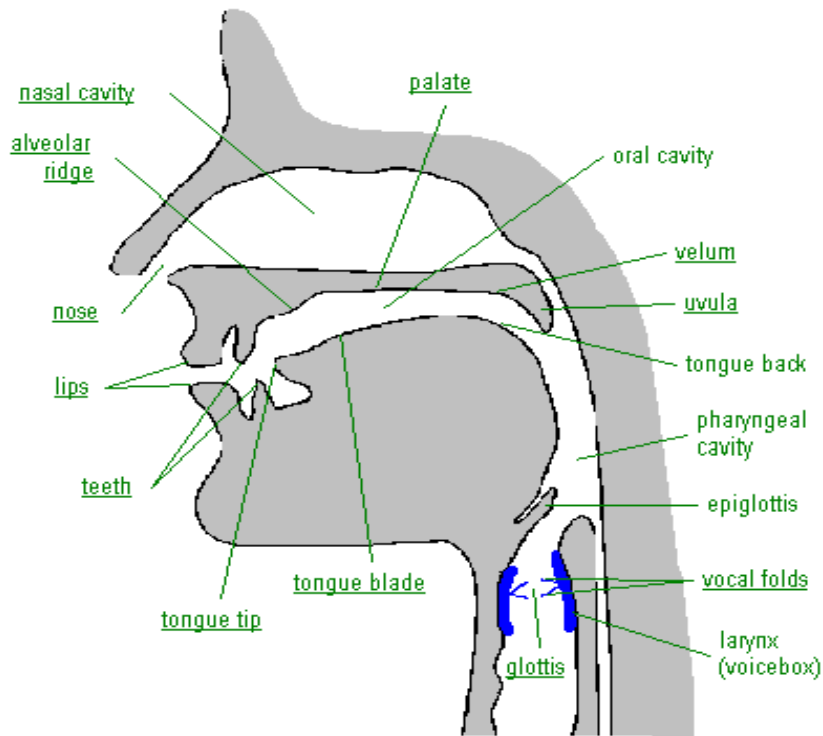
The lesson plan was incorporated into the Articulatory training session using a whiteboard. *pill*, *spill*, *hill* and International Phonetic Alphabet (IPA) symbols were written on the board, and were pointed to—to avoid articulating the sounds myself—while the subjects mimicked the sounds. Subjects in this group were allowed to put their hands in front of their mouths to feel aspiration. I indicated to both groups to pay close attention to whatever differences they could perceive between any and all sounds in the training session. Next, I wrote *s...pill*, *s...tan*, and *s...school* to progress to plain voiceless stops in word-initial position. I drew a sagittal diagram of the vocal tract for the

Articulatory group to show place of articulation. To demonstrate word-medial and word-final production of voiceless aspirated Hindi consonants, I transliterated the words from part 4 of the lesson plan above: *seekhna*, “*pitch*”*lay*, *goop^haye*, *aangkh*.

When I introduced dental and retroflex stops to the Articulatory group, I was able to show them place of articulation on the sagittal diagram. I showed them that the tongue pushes against the back of the upper teeth when making a dental sound, and curls back toward the hard palate to make a retroflex sound. I avoided terms like “hard palate” to make the lesson less complicated; instead, I drew pictures for what each sound production looked like.

The Auditory group did not get to see this. Instead, I played them sound files of dental and retroflex consonants, with IPA symbols written on the board for their reference. Dental and retroflex sounds are not recognized as unique phonemes in English, but sometimes occur in rapid speech. (For example, I pronounce “in there” as [ɪ̠nd̪ɛɪ] because my tongue combines two places of articulation while moving quickly.)

Those subjects in both training groups who were Linguistics students *did* have prior knowledge of these terms, and therefore knew what types of sounds they were supposed to produce. Nevertheless, their familiarity with the terms did not give them superior skills in production; they, like non-Linguistics subjects, experienced difficulty in producing the sounds.



Copied with permission from Instituto Lingüístico de Verano, A.C.

The diagram above is a sagittal view of the vocal tract, similar to the one I drew for the Articulatory group. The sounds that were taught in the training sessions are all concentrated in the oral cavity (from the lips to the uvula).

Both training groups had a chance to compare alveolar [t], dental [t̪] and retroflex [ɭ]. With Articulatory trainees, I used the sagittal diagram to point to the changing positions of the tongue. I demonstrated how the dental [t̪] is where the tongue hits the back of the upper teeth, the alveolar [t] touches at the alveolar ridge, and the retroflex [ɭ] hits the hard palate (indicated by “palate” on this diagram). This holds true for their voiced counterparts as well.

Finally, I moved to voiced aspiration with the Articulatory group. As with the Auditory group, I told them that we would try to combine [b, ɓ, ɗ, g] with *hill* to get a voiced aspirate in word-initial position. I had them repeat each stage of the progression

from part 6 of the lesson plan above: starting with *b...hill* and ending with [b^hɪl]. I made the subjects in both groups practice multiple times, and I gave feedback such as “good, you’re getting closer” or “okay, try to fuse the sounds a little more”.

My last step with both training groups was a quick review of what they had learned. To add some humor to the session, I had them try to pronounce my last name—[b^hʌt]—as proof that the training methods had worked. A few subjects from the Articulatory training group pronounced it correctly.

5. Procedure

All testing and training was held in the UC Berkeley Phonology Laboratory. Subjects attended the pre-test in groups of one to six. They were introduced to the study and provided with testing forms. The forms were based on the same-different paradigm. Thirty pairs of sounds were presented: each pair was played three times, and after the third repetition the subjects circled “same” or “different”. This method employs a “delayed response” similar to the “delayed shadowing” expressed in the study on visible speech and perception (Davis & Kim, 2001). The test took ten minutes.

Following the pre-test, I set up times and dates for the training sessions through e-mail.

The training sessions took place about two weeks after the preliminary test for five of the non-Hindi learners, and about one or two days after the preliminary test for the other five. This was based entirely on schedule coordination, and did not have any impact on results.

I held both training sessions on the same day for those five subjects who came two weeks later. I instructed both sessions. Two of the subjects attended the Articulatory

training session in the morning, and three attended the Auditory training session in the afternoon.

For the remaining five subjects, the Articulatory training session was held on a separate day from the Auditory training session—again, only due to scheduling. This had no impact on the results.

Five Group A subjects took the production exam immediately after the post-test. The remaining five returned after one or two days. This was once again due to scheduling conflicts.

The subjects in Group B—current or former Hindi learners—were not required to attend any training.

During the post-test, subjects were provided with the same forms used in the preliminary test, and were explained the same rules. The post-test took ten minutes.

The individual production exam was conducted in the Phonology Laboratory sound booth (52A Dwinelle Hall). Each test lasted between five and eight minutes.

I conducted the first production exam with A.H. using Wavesurfer v. 1.8.5, but the quality came out grainy. I continued with Audacity v. 1.2.4 (Mazzoni, Brubeck, Crook, Johnson, & Meyer, 2006) for all of the following production exams. When I played A.H's sound file again after completion of the study, the quality was fine.

Each subject sat in front of the computer in the sound booth and wore a microphone that hooked behind the ears. I sat next to the subject, without a microphone. I read from a pre-written list of thirty made-up words that had each of the ten Hindi aspirated consonants in word-initial, word-medial, and word-final position. (Please refer to Section 3 for the list of words.) I said each word, and the subject immediately

mimicked me while recording onto a sound file. Since it was difficult to expand the waveform while recording to show the subject his/her articulation, the subject had to rely primarily on auditory perception.

Upon completion of the production exam, the subject signed a receipt form and received payment. I paid them either the full \$15 or the remaining amount that they had not yet received.

Testing was completed on Friday, February 29th, 2008 for fourteen subjects. One subject was unable to complete the production exam.

6. Results and Discussion

Results from my study differed from those of Catford and Pisoni's study, for various reasons. Two main reasons include my changes to the experiment: first, I added the Hindi learners as a control group, since they had had prior exposure to the language of focus. Second, I created the preliminary test, which established each subject's level of perception of Hindi consonants before the training session. The pre-test/post-test design, rather than training and a post-test as was done by Catford and Pisoni, helped to measure *improvement* in perception.

One thing that did not change between the pre- and post-test was the Hindi learners' superior ability in both discrimination and production. This result was expected. Prior L2 exposure to the language of focus provided the Hindi learners with two advantages: 1) accurate recognition of how each Hindi aspirated consonant is supposed to sound; and 2) accurate or near-accurate production of each consonant. Naturally, the Hindi learners obtained the highest scores in both perception and production. This proves the hypothesis that prior L2 exposure to the language of focus provides an advantage over

systematic instruction (Au, Knightly, Jun, & Oh 2002; Werker & Logan 1985; Chomsky 1981). This holds true even over an extended period of time, considering some Group B students had not studied Hindi for several semesters.

I expected to verify Catford and Pisoni's results regarding the training groups; that is, I expected the Articulatory training group to surpass the Auditory training group in both discrimination and production. However, my results turned out differently.

The Auditory training group scored more accurately on the post-test discrimination task than the Articulatory training group. Both the pre- and post-test consisted of twenty "same" and ten "different" responses. Total "same" responses numbered 300 (20 "same" responses per test x 15 subjects = 300 total "same" responses), and total "different" responses numbered 150 (10 "different" responses per test x 15 subjects = 150 total "different" responses). Only 3 out of 150 total "different" responses were marked incorrectly as "same". The real source of error was in marking too many "same" answers as "different"; that is, *over-discriminating* between sounds. It made for a more interesting and fruitful analysis to concentrate on this error.

Thus, I scored each subject on his/her percent-accuracy of "same" responses out of twenty (20 total "same" responses per test).

The chart below indicates the average percent-accuracy on the pre-test and the post-test for the Articulatory group, the Auditory group, and the Hindi learners.

(Chart 2.1 in the Appendices contains the same data as the table below. Charts 2.2-2.4 display individual subjects' percent-accuracy.)

Average % accuracy on Discrimination

	Pre-test	Post-test
Articulatory	71.25	70
Auditory	75	78.33
Hindi	98	95

According to the chart, the Auditory group surpassed the Articulatory group on both the pre-test and the post-test. However, the pre-test occurred before training took place. (I only sorted the pre-test data into training groups to make the chart easier for the reader.) Thus, the superior pre-test score of the Auditory group over the Articulatory group is *not* attributable to the Auditory training method, but their superior post-test score *is*.

Rather, the pre-test score is attributable to certain individuals' higher sensitivity to Hindi consonant discrimination. This came as a surprise, since I was only focusing on the expected gap in scores between Hindi learners and non-Hindi learners. The heightened perception of some non-Hindi learners is most likely a result of prior exposure to similar sounds in other languages.

(Please refer to Appendices, Table 1 for each subject's language background.)

For example, B.A, J.M., and L.C. are the three Group A subjects who obtained 75% or higher on the pre-test and who were all assigned to the Auditory group for training. B.A. and J.M. have had regular exposure to French at UC Berkeley. French contains the voiced dental [d̪]. L.C. has had regular exposure to Taiwanese at home—which contains the voiced affricate [dʒ]—and studied Spanish in high school, a language which also contains the voiced dental [d̪]. This prior L2 exposure is the most probable reason why these subjects did better than others on the preliminary discrimination task.

Since I divided subjects at random into training groups, it was a coincidence that B.A., J.M., and L.C. were all assigned to the Auditory group. I did, however, purposely place more subjects in the Auditory group (6) than the Articulatory group (4). I did this to challenge the results obtained by Catford and Pisoni using the Articulatory training method. If the Articulatory method was indeed superior to the Auditory method, the Articulatory group should have obtained higher scores than the Auditory group even if it had a smaller subject population.

Apparently, this did not hold true. The Auditory group still did better on the discrimination task. Despite incorporating the majority of Catford's teaching method into my two training sessions, I did not replicate Catford and Pisoni's results.

I can identify two reasons why this happened: first, as explained above, the Auditory group's superior pre-test score was not a result of Auditory training, but heightened individual sensitivity that was unrelated to this experiment and that must have been a result of previous L2 exposure. Once these subjects were randomly placed in the same training group, their combined sensitivity skewed the data and gave the Auditory group a significant perceptual advantage over the Articulatory group.

Second, I improved the Auditory training session by providing the group with audio playback of sound files of Hindi breathy consonants. I did this to balance the two training sessions, because the Articulatory training group also had auditory exposure through their own mimicry of sounds. I wanted to refrain from producing sounds in either training session, so I decided that the Auditory group needed an added source of auditory exposure. I did not, however, change Catford's method of teaching in either session; I simply changed the lesson plan.

