Acquisition of Stops of Hindi by Speakers of American English

Lauren Bonomo

Advisor: Jelena Krivokapić

Yale University, Department of Linguistics

May 1st, 2013
Table of Contents

Abstract........................................................................................................................................3
1. Introduction....................................................................................................................................4
2. Background.....................................................................................................................................4
  2.1 The Critical Period Hypothesis.................................................................4
  2.2 Other Factors Influencing Second Language Acquisition...............5
  2.3 Speech Learning Model.........................................................................................6
  2.4 Languages in the Present Study.................................................................7
    2.4.1 American English....................................................................................8
    2.4.2 Hindi........................................................................................................9
  2.5 Research Question and Hypothesis..........................................................10
3. Experiment..............................................................................................................................11
  3.2 Methods.........................................................................................................................11
    3.2.1 Participants..............................................................................................11
    3.2.2 Data Collection..........................................................................................12
    3.2.3 Stimuli.........................................................................................................12
    3.2.4 Data Labeling............................................................................................13
    3.2.5 Data Analysis.............................................................................................14
4. Results and Analysis..............................................................................................................15
  4.1 Native Group..............................................................................................................15
  4.2 Advanced Group.......................................................................................................16
  4.3 Beginner Group.......................................................................................................16
  4.4 One Beginner Subject Omitted..................................................................................17
5. General Discussion................................................................................................................19
6. Conclusions..........................................................................................................................21
References......................................................................................................................................22
Appendix: Stimuli.........................................................................................................................24
Acknowledgements..................................................................................................................25
Abstract

The extent of the validity of the Critical Period Hypothesis (CPH) has been one of the central debates in second language acquisition for decades. One theory that contradicts the CPH is Flege’s Speech Learning Model (Flege 2005). This model proposes that we maintain the tools necessary to acquire a new language with native-like proficiency throughout our lives. Flege’s theory attributes the variation in success in second language acquisition to the quantity and quality of input in addition to the identity of a speaker’s first language, rather than biological changes. He proposes that the relative rates at which learners will acquire the phonetic properties of a new language can be predicted based on these factors. The present study aims to contribute to the dialogue surrounding the CPH by testing the hypotheses generated by Flege’s model, specifically the hypothesis that sounds in a second language more dissimilar to the closest sound in one’s native language will be learned more readily than extremely similar sounds. Production of four Hindi stops was examined in two groups of native English speakers learning Hindi for one year or two years, along with a control group of native Hindi speakers. Results show a greater ease of learning for the Hindi phonemes most dissimilar from the closest in English, lending support to the Speech Learning Model and suggesting the need for modification of the Critical Period Hypothesis.
1. Introduction

Even those unfamiliar with the field of linguistics have several basic observations about language acquisition. First, almost every child acquires a first language (L1) to native proficiency. Unlike this nearly universal success in L1 acquisition, learning a second language (L2) has much more varied results (Birdsong 2006). In linguistics, second language acquisition is an area of vigorous research, approached by researchers from a wide range of linguistic disciplines. These include, but are not limited to, linguistic theory, evolution theory, language processing, and neurophysiology (Birdsong 1999). The present study aims to add to the discussion, focusing on phonetic acquisition of a second language in adult learners.

2. Background

2.1 The Critical Period Hypothesis

For much of the late twentieth century, the most widely accepted theory regarding L2 acquisition was built on the observations listed above. This theory, called the Critical Period Hypothesis, was introduced by Lenneberg (1967) and supported by several early studies (Krashen 1973, Snow and Hoefnagel-Hohle 1978). It states that, “there is a limited developmental period during which it is possible to acquire a language to normal, native-like levels” (Birdsong 1999). After this window, capacity to learn a language declines, and native-like proficiency is seldom achieved. This hypothesis was accepted for years as the main factor influencing second language acquisition (Johnson and Newport 1989).
2.2 Other Factors Influencing Second Language Acquisition

Following the work of Lenneberg and his supporters, theories began to emerge that challenged the strong view of the CPH. New data suggested that there were many more factors influencing the process of second language acquisition than merely age. For instance, it was shown that early arrival in a country was positively correlated with attainment of native norms, even if arrival was past puberty (Birdsong 1992, Flege et al. 1995, MacKay 2001). Several researchers observed native-like accuracy in their late-learning subjects as well (Mayberry 1993, Van Wuijtswinkel 1994, White and Genesee 1996). One asserted that the disadvantages of a late start could be overridden entirely to produce speech with no foreign accent at all (Bongaerts et al. 1997).

In addition to showing the high degree of potential for certain late-learning subjects, studies also showed that early learning was not always linked to success. Flege discovered that a group of Korean children in Canada had the same mean foreign accent rating both 3.5 and 5.5 years after arrival (Flege et al. 2006). The same study showed that 10-year-old Koreans who arrived in Canada at the age of six had significantly lower ratings than 10-year-old native speakers. Flege also found that the degree to which two groups of Canadian residents, whose L1s were Italian and Korean, spoke with an accent was more strongly linked to identity of the L1 than to age of arrival in Canada (Flege et al. 1999). These studies laid the groundwork for the development by the author of a model of L2 acquisition that does not depend on age alone.
2.3 Speech Learning Model

Flege names his proposed model of second language acquisition the Speech Learning Model. According to Flege, “there is no conclusive support for the existence of a critical period in human speech learning” (Flege 1987). His theory maintains that the processes and mechanisms that make first language acquisition possible remain intact and accessible throughout a person’s lifetime. Importantly, these mechanisms include the ability to form new phonetic categories, given sufficient native speaker input (Flege and Liu 2001). The Speech Learning Model proposes that the way new sounds will be acquired and the degree of success in L2 production is heavily dependent on L1.

One way that L1 can affect second language acquisition is through allophonic interference. In one experiment, Flege showed that English speakers learning French were more readily able to learn the French /y/ than the French /u/ (Flege 1984). Both French sounds are heard as the English /u/, but /y/ is significantly more fronted than the other two. The French /u/ is much more similar to the English sound. Several of the testable predictions generated by the Speech Learning Model have to do with this phenomenon.

First, the greater the dissimilarity perceived in an L2 sound from the most similar L1 sound, the more readily a new phonetic category will be formed for that sound. Second, when a new category is established for an L2 sound, it may dissimilate from neighboring sounds to preserve phonetic contrast. Finally, when a unique category is not formed for an L2 sound because it is too similar to its L1 counterpart, the categories for L1 and L2 will assimilate, leading to a merged L1-L2
value (Flege and Hammond 1982). Data from Flege’s experiment on English speakers learning French provided support for these hypotheses in vowels. The present study aims to do so in consonants.

Flege’s most similar work involving consonants is a study of the phonetic implementation of the stop voicing contrast produced in Arabic by Saudi Arabians and by both Americans and Saudi’s in English (Flege and Port 1981). This study demonstrated that, although Arabic has voice and place features for stops, subjects had a greater difficulty producing the voiced-voiceless distinction at a place where the distinction does not exist in Arabic. The study was useful in demonstrating that this sort of problem in L2 speech learning is phonetic rather than phonological, but leaves the aforementioned hypotheses of the Speech Learning Model to be examined in consonants.

2.4 Languages in the Present Study

Conducting this study at an American university, the majority of possible test subjects are native speakers of American English. It was critical that the experimental L2 contain consonants that differed from those of American English in a way that could be empirically measured. This led to the choice of Hindi as the experimental L2, as its consonants differ from those of English in voice onset time (VOT), a quantifiable characteristic.
2.4.1 American English

In English, word-initial voiceless stops are aspirated. When the same phonological units appear mid-word or word-final, however, they are typically unaspirated. A more complete statement of the phonological rule describing the phenomenon is that a voiceless stop that does not associate with the preceding syllable but is syllable-initial becomes aspirated. Therefore, word-initial voiceless stops in English have a measurably longer VOT than the same word-final, mid-syllable, or unstressed mid-word stops. For instance, the VOT of the phoneme /t/ in the English word “tie” is markedly longer than that of the same phoneme in the word “sty.” These otherwise identical phonemes of distinct VOT are not contrastive in English. This phenomenon is not present in voiced stops.
Figure 1. Example waveform of the three alveolar stops of English (Ladefoged 2010). The first panel shows aspiration in the word-initial voiceless stop, with the highest VOT of the three. The third panel shows the same phoneme mid-syllable, exhibiting the shortest VOT of all. The VOT of the voiced stop, seen in the second panel, is only slightly longer than that of the mid-syllable voiceless stop.

2.4.2 Hindi

Hindi, unlike English, has contrastive aspirated and unaspirated stops. All voiceless stops and most voiced stops in Hindi have aspirated and unaspirated contrastive options (Ladefoged 2010). Ladefoged presents typical VOT data on the two stops examined in this experiment in the following figure.
10

Figure 2. Example waveform of the four dental stops of Hindi (Ladefoged 2010). The three topmost waveforms are extremely similar to the typical waveforms of the three alveolar phonemes of English seen in Figure 1. They correspond to English’s voice stop, word-initial voiceless stop, and non-stressed mid-word or mid-syllable voiceless stop, respectively. The fourth panel shows typical onset of vowel for the aspirated, voiced stop of Hindi, which has no English equivalent.

2.5 Research Question and Hypothesis

According to the CPH, native speakers of English should be successful in reproducing those three phonemes in Hindi (/t/, /th/, and /d/) and should generally fail to produce the fourth, which has no corresponding English sound.

This study aims to advance Flege’s model through an examination of native speakers of American English learning Hindi. If the model is indeed an accurate predictor of patterns in L2 phonetic acquisition, subjects should easily create a new phonetic category for Hindi’s aspirated /d/, which has a much higher VOT than the
corresponding voiced stop of English. Early learners may even overcompensate and delay onset of the vowel further than native speakers. Native English speakers should have more difficulty acquiring the aspirated and unaspirated /t/, for highly similar sounds exist in English but are not contrastive. I also hypothesize that, over time, learners will begin to form distinct phonetic categories for aspirated and unaspirated /t/ as well (Flege et al. 1999). Findings consistent with these hypotheses would support Flege’s Speech Learning Model and add evidence to the theory that people maintain the mechanisms and processes that guide language acquisition throughout their lifespan.

3. Experiment

This present study examines the production of the Hindi aspirated-unaspirated contrast in two stops, the dental /t/ and /d/. It aims to contribute to the resolution of some of the debate surrounding the CPH of second language acquisition by assessing one model of phonetic acquisition. Data was collected from learners at various stages of acquisition, including native speakers of the target L2 as a control.

3.2 Methods

3.2.1 Participants

Twelve subjects whose ages ranged from eighteen to twenty-four took part in the study. These subjects were divided into three groups: native Hindi speakers, native English speakers learning Hindi for less than one year, and native English speakers learning Hindi for approximately two years. These groups are referred to in the
following charts and discussion as “Native,” “Beginner,” and “Advanced,” respectively. Each of these groups contained four subjects. Subjects were not informed of the purpose of the experiment until after they had completed the tasks.

3.2.2 Data Collection

All subjects were recorded in the sound booth at Yale University with a Logitech Noise-Cancelling USB Desktop Microphone. Recordings were done with the MacOS X edition of Praat (Boersma and Weenink 2013). Subjects were told to read a series of sentences as if they were reading a story. Twenty-four sentences were recorded at a time and saved into an online backup service, named with the subject’s level and a letter indicating the order in which they participated. For example, data from the first Native speaker recorded would be saved as “Native_A.” Five consecutive recordings were made for each subject.

3.2.3 Stimuli

The elicitation materials for the data reported in this paper were presented in one list of 120 Hindi phrases. Each phrase was of the same form -

फिर से “X” बोलो।

*Phir se “X” bolo.*

Say “X” again.

- where X was either a token or filler word. There were three sets of four token words and twelve filler words. Both token and filler were monosyllabic, real Hindi words of the form CVC. The three sets of four token words differed in their central
vowel (/æ/, /ə/, and /ʊ/) and contained one word with each of the following onsets: /t/, /tʰ/, /d/, and /dʰ/, allowing the acquisition of the four stops to be tested. For example, the first set of words was:

ताल - थाल - दाल - धार
taal – thaal – daal – dhaar

The list of filler material contained an equal number of monosyllabic, real Hindi words of the form CVC. These twenty-four phrases were randomized five times for a total of sixty data points per subject.

### 3.2.4 Data Labeling

Data were labeled using the MacOS X edition of Praat (Boersma and Weenink 2013), as shown in Figure 3. First, all recordings were listened to and tokens were marked using a point tier in TextGrid. Subsequently, release of the stop closure and onset of voicing were also marked. The time between the release of the stop closure and the onset of voicing (VOT) was calculated. There were four cases for which the onset of voicing could not accurately be determined, and those tokens were omitted from data analysis.
Figure 3. Sample data labeling. The release of the stop and onset of voicing were both marked and labeled in a point tier, and the length of time between them was visible in the grey bar below. This measurement, VOT, was recorded in seconds. This waveform comes from a Native subject saying the word *thaal*.

### 3.2.5 Data Analysis

VOT measurements were first standardized (Z-scored) to render data comparable across individuals. This was accomplished using the IMB SPSS Statistics Base 21 for Mac Descriptive Statistics “Descriptives” function. After Z-scoring, a repeated measures analysis of variance (ANOVA) was performed in SPSS. In the analysis, the dependent variable was VOT, and the independent variable was the identity of the phoneme. Significance was accepted at p < 0.05. Additional script was written to generate a pairwise comparisons table after post-hoc tests showed significance, so that the difference between any two phonemes within a single group was readily visible.
4. Results and Analysis

The repeated measures ANOVA showed significant within-subject contrasts among the four phonemes ($F(3) = 109.481, p = 0.000$) and significant phoneme contrasts among the three groups ($F(6) = 6.756, p = 0.000$). These results prompted the differences in standardized VOT to be more closely examined within the subject groups.

4.1 Native Group

Any trends noticed in the Beginner and Advanced learner groups would not be meaningful without being compared to trends in Native speakers of the target language. It would not be useful to know whether learners’ VOT for the phonemes in question were significantly different if this were not the case for native Hindi speakers. Post-hoc pairwise comparisons showed statistically similar VOT for the two unaspirated stops ($p = 0.356$) and for the two aspirated stops ($p = 0.445$). The difference between each aspirated stop and its unaspirated counterpart was, however, significant ($p = 0.000$ for both pairs). There was thus a marked difference in standardized VOT between the aspirated and unaspirated version of each stop. This finding indicates that VOT will indeed be a useful measure of how accurately learners pronounce the aspirated-unaspirated distinction in Hindi stops and whether new phonetic categories are established.
4.2 Advanced Group

The next group was the Advanced group, the subjects who had been learning for roughly two years. In this group, the same effect that occurred in the Native speaker group was observed. The VOT was not significantly different between the two unaspirated stops (0.654) and two aspirated stops (0.280), but it was significantly different between each aspirated-unaspirated pair (p = 0.000 for both). These results indicate that, although the test also showed that the VOT of each individual phoneme had a greater variance than those of native speakers, the aspirated and unaspirated versions of each sound had statistically different VOTs. These data indicate that the Advanced speakers have different phonetic categories for all four Hindi phonemes.

4.3 Beginner Group

The final group was the Beginner group, whose subjects had been learning for less than one year. This group, like the Native and Advanced groups, showed significantly different standardized VOTs for the aspirated and unaspirated voiced stop (p = 0.034). Unlike the first two groups, the Beginner group did not show statistical difference between the aspirated and unaspirated voiceless stop (p = 0.725). These data indicate that the two /d/ sounds have significantly different VOTs in this group of early learners; the difference between the /t/ sounds, on the other hand, is not significant.
Figure 4. Estimated marginal means of standardized VOT for each phoneme in each subject group. It is easily observed that the earliest learners mastered the aspirated-unaspirated distinction in the voiced stop more readily than they did for the voiceless stop. The distance between aspirated and unaspirated VOT for the voiced stop is greatest in the Beginner group, while that of the voiceless stop is the least by far.

4.4 One Beginner Subject Omitted

Finally, a figure was created with the exclusion of one of the beginner subjects. This subject (“Beginner Subject B”) had been exposed to Hindi for less than one year, but, in addition to the beginner language course, he also participated in Hindustani classical vocal training. So while the duration of his input was consistent with the other subjects in the beginner group, the quality of the input was enhanced. If
Flege’s claims about quality of input are accurate, removal of this Subject B’s data from the figure should result in the trends in Figure 4 becoming more pronounced. Figure 5 shows that this is indeed the case. Results of a repeated measures ANOVA excluding data from Subject B did not change in significance.

Figure 5. Estimated marginal means of standardized VOT excluding Beginner Subject B. It is observed that without data from the Beginner subject with the best quality (and perhaps greatest quantity) of input, the mean VOTs for the aspirated and unaspirated voiceless stops are even more similar.
5. General Discussion

To reiterate, Flege’s Speech Learning Model generates three testable hypotheses. First, we will discuss the hypothesis directly addressed by the present study: that the more dissimilar an L2 sound is from its closest L1 sound, the more readily a new phonetic category will be formed. The data obtained in this experiment strongly support this hypothesis. Even the earliest learners were shown to pronounce the aspirated /d/, which does not exist in English, in a manner distinct from its unaspirated counterpart. The seemingly easy acquisition of the aspirated /d/ of Hindi does not support the strong view of the Critical Period Hypothesis. The pair of /t/ sounds, however, presented a more difficult case for native English speakers, whose language contains non-contrastive sounds very similar to both. It can be seen in Figure 4, and more so in Figure 5, that the production of the aspirated and unaspirated voiceless stops by early learners is not significantly different.

While Flege’s model predicts greater difficulty creating new categories for phonemes similar to L1 sounds, it also predicts that doing this is possible in time. The data obtained from the Advanced group shows this happening. As subjects became more proficient in Hindi, their aspirated-unaspirated contrast became significant. The means of standardized VOT of the Advanced group, shown in Figures 4 and 5, are quite similar to those of the Native group. The greater variance observed in the Advanced group suggests the phonetic categories for these phonemes might not be as well established as they are in native speakers, but they are a marked improvement over the means of the Beginner group.
The Speech Learning Model's next hypothesis is that sounds in a newly established phonetic category may dissociate from neighboring sounds to preserve phonetic contrast. This is also consistent with data obtained in this experiment. Figures 4 and 5 show the greatest difference in VOT between the aspirated and unaspirated voiced stop was observed in the Beginner group. The next greatest difference was observed in the Advanced group, with the smallest difference in the Native group. The newly established phonetic category was observed to be more dissociated from its unaspirated counterpart the earlier in the process of acquisition the group was. This is likely the result of an effort to over-distinguish the difference between the already-familiar unaspirated stop and the new aspirated stop.

Flege’s final hypothesis was not directly addressed in the experimental design but may have in fact played a role in the data obtained. This hypothesis is that when new categories are not formed because the L2 counterpart is too similar, as in the case of unaspirated /d/, which is quite similar in English and Hindi, categories for both sounds may assimilate. According to data presented by Ladefoged and shown here in Figure 1 (Ladefoged 2010), the Hindi voiced unaspirated stop has a VOT near or less than zero. This was not observed in this study’s native speaker data, with every utterance of the sound having a positive VOT and an aggregate VOT almost equal to that of the unvoiced aspirated stop. This could be a case of convergence, the result of all the native Hindi speakers participating in this study also speaking English fluently. Studies have shown that phonetic production can converge when speakers are exposed to a new linguistic environment (Pardo et al. 2012). It is possible that long-term exposure to American
English as the ambient language led to lengthening of VOT in the unaspirated /d/ of Hindi, resulting in production more similar to the /d/ of English. If this were the case, it would lend further support to the Speech Learning Model, as this sort of assimilation between extremely similar phonemes is predicted by the model.

6. Conclusions
While not providing a general conclusion regarding the Critical Period Hypothesis, the central thrust of which may be valid, this study supports modification by incorporating Flege’s Speech Learning Model. The hypothesis tested by this study, that new phonetic categories will most readily be formed for sounds with the greatest dissimilarity from the most similar L1 sound, was strongly supported. Additionally, several other hypothesis generated by the model that were not directly tested seem to have been supported as well. The Speech Learning Model’s ability to predict how learners will acquire new sounds in an L2 suggests that the identity of L1 is a more significant influencing factor than age. Based on the data obtained in this study, it can be said that second language acquisition is a more complicated process than the strong view of the Critical Period Hypothesis would indicate. Biological change may not be the most significant factor influencing ultimate attainment in a second language.
References


Apendix: Stimuli

Token “X”:
फिर से “तल” बोलो।
फिर से “थल” बोलो।
फिर से “बल” बोलो।
फिर से “घर” बोलो।
फिर से “तुक” बोलो।
फिर से “थुल” बोलो।
फिर से “दुख” बोलो।
फिर से “धुल” बोलो।
फिर से “ताल” बोलो।
फिर से “थाल” बोलो।
फिर से “दाल” बोलो।
फिर से “धार” बोलो।

Filler “X”:
फिर से “जल” बोलो।
फिर से “भाई” बोलो।
फिर से “बस” बोलो।
फिर से “मर” बोलो।
फिर से “चार” बोलो।
फिर से “पर” बोलो।
फिर से “मार” बोलो।
फिर से “तीन” बोलो।
फिर से “सात” बोलो।
फिर से “मन” बोलो।
फिर से “चल” बोलो।
फिर से “नान” बोलो।
Acknowledgements

I would first and foremost like to thank Professor Jelena Krivokapić for her help and understanding throughout this project. Although my study got a somewhat late start, she was never discouraging, and her expertise, standards, and patience made completing it possible. Without her ideas and support, the task of the senior essay would have been a much more daunting one.

I would also like to sincerely thank Professor Raffaella Zanuttini, not only for her guidance relating to the senior essay, but for four years of help and advice as well. She was one of the first professors I had at Yale, and her excellence in teaching and passion for linguistics are a large part of what inspired me to pursue this major. Without her recommendation, I would not have been admitted to medical school early enough to fully pursue my passion for language learning, and for that I am deeply grateful.

Finally, I would like to thank the other linguistics majors in the Class of 2013. The incredible knowledge, talent, humility, and sense of humor possessed by this extraordinary group enriched my time at Yale and made the senior seminars not just manageable, but delightful.