Stop Consonant Characteristics: VOT and Voicing in American-Born-Indonesian Children’s Stop Consonants

A Senior Honor Thesis

Presented in Partial Fulfillment of the Requirements for Graduation

“with Honors Research Distinction in Speech and Hearing Science” in The Speech and Hearing Science Department of The Ohio State University

by:

Florence Lim Hardjono

The Ohio State University July 2011

Research Advisor: Robert A. Fox, Ph.D, Department of Speech and Hearing Science
For Jeremy, Sammy, and Gracie. My world changed for the better when you were born. Your language and cognitive development never cease to amaze me. May God allow me to carefully learn from you and other bilingual children throughout my academic journey. This is only the beginning.

Glory be to God the creator of language, who painted the sky and designed the amazing wonders of human beings.
Acknowledgments

I would like to thank my research advisor, Dr. Robert A. Fox, for his continuous guidance and support throughout this research project. Dr. Michelle S. Bourgeois, my academic honors advisor for her guidance and support throughout the writing process. Dr. Ewa Jacewicz, co-director of Speech Perception & Acoustics Laboratories for her support and guidance throughout the recording and data analysis process. Dr. John Grinstead, thank you for your willingness to be a part of the committee. Dr. Mary E. Beckman from the Ohio State University Linguistics Department for allowing me to assist Elizabeth A. McCullough (her doctoral student) in her bilingual VOT research. That initial exposure to research ignited my research interest in bilingual children language acquisition.

I am grateful and thankful for my subjects who took their time to participate in this project. In addition, I am indebted and grateful for my friends Helen Santoso, Maya Dewi, Natalie Gunawan, Irma Khouw, Ginger Wahjudi and Jean Johnson, who so willingly took care of my kids whenever I need to go to OSU campus.

Most importantly, I would not be able to complete this project without the loving support of my husband, Luki Hardjono. Thank you for holding of the battle at home and yes that includes all those “laundries” you washed and “meals” you microwaved.

This project was partially funded by an ASC scholarship, an SBS scholarship and Speech Perception & Acoustics Laboratories. My gratitude and thanks to the Ohio State University for its continuous support for research.

Finally, to my God who said, “For I know the plans I have for you,” declares the LORD, “plans to prosper you and not to harm you, plans to give you hope and a future.” Jeremiah 29:11.
Table of Contents

Abstract .............................................................................................................................................. 6
List of Tables ........................................................................................................................................ 7
List of Figures ....................................................................................................................................... 8
Chapter 1: Introduction .......................................................................................................................... 9
Chapter 2: Literature Review ................................................................................................................ 11
  VOT ................................................................................................................................................... 11
  Voicing ............................................................................................................................................. 12
Chapter 3: Methodology ........................................................................................................................ 14
  Subjects ............................................................................................................................................. 14
  Preliminary procedure ....................................................................................................................... 15
  Stimuli ............................................................................................................................................... 16
  Data collection / Recording ............................................................................................................... 17
  Data analysis ................................................................................................................................... 18
Chapter 4: Result .................................................................................................................................... 23
  Voicing ............................................................................................................................................. 23
  Mean VOT and closure duration across speaker types ................................................................. 24
  Mean VOT for Indonesian adults based on home language preference ....................................... 27
  Mean closure duration for Indonesian adults based on home language preferences ............... 27
  Mean VOT for Indonesian children based on home language preference ................................ 28
  Mean closure duration for Indonesian children based on home language preferences ............ 28
Chapter 5: Discussion .......................................................................................................................... 30
Chapter 6: Summary/Conclusion .......................................................................................................... 33
Abstract

The increasing number of Indonesian immigrants in the United States suggests the need for deeper understanding of English acquisition as a second language by Indonesian children and their adult caregivers. Unfortunately this language’s interaction with English as an L2 (particularly in terms of language interference) is currently understudied. This study examines the acquisition of English stop consonants by typically developing Indonesian bilingual children and their parents. Thirteen pairs of parent-child Indonesian bilingual speakers produced six stop consonants (/p, t, k, b, d and g/) common in both languages in word-initial position with a variety of vowels. Pictures representing these words were shown on a computer monitor and speakers were asked to say the word in the context “That’s a ___” for English words and its equivalent “Itu ___” for Bahasa Indonesia words. The phonetic characteristics of voice onset time were the focus of this study. The prediction is that the children’s production of English stops from both subject groups will have similar acoustical characteristics. However, their parents’ production will have different acoustical characteristics because of the influence of the age of learning a second language. The findings of this study will enable us to learn more about first and second language acquisition among bilingual children.
## List of Tables

Table 1. Subjects’ gender, birth country and home language preferences. ........................................ 14
Table 2. Children’s ages .................................................................................................................... 15
Table 3. Result of voicing categories across speakers. .................................................................. 23
Table 4. Mean VOT across speaker types. ....................................................................................... 24
Table 5. Closure duration across speaker types for voiced and voiceless stops. ......................... 26
Table 6. Mean VOT based on home language preferences for Indonesian adults. ..................... 27
Table 7. Mean closure duration based on home language preferences for Indonesian adults. ..... 27
Table 8. Mean VOT based on home language preferences for voiced and voiceless stop consonant for Indonesian children. ..................................................................................... 28
Table 9. Mean closure duration based on home language preferences for voiced and voiceless stop consonants for Indonesian children. ........................................................................... 29
List of Figures

Figure 1. Sample of an English stimulus ................................................................. 17
Figure 2. Markers for calculating VOT and closure duration ........................................ 19
Figure 3. An example of no voicing on a spectrogram .................................................. 20
Figure 4. An example of pre-voicing on a spectrogram .................................................. 20
Figure 5. An example of complete voicing on a spectrogram .......................................... 21
Figure 6. An example of partial voicing on a spectrogram ............................................. 21
Figure 7. Mean VOT across speaker types .................................................................... 25
Figure 8. Mean closure duration across speaker types .................................................. 26
Figure 9. Mean VOT for each consonant across speaker types ....................................... 30
Chapter 1: Introduction

The increasing number of bilingual children in the United States of America requires careful examination that looks into their language acquisition. The native language (L1) one learns in early childhood and a second language (L2) learned later in life often influence one another (Flege, 1991). Therefore, it is important for speech language pathologists to be aware of these differences in phonetic characteristic of bilingual children. Adequate knowledge about speech development in normal developing children from culturally and linguistically diverse background will enable speech language pathologists to work more effectively.

Currently, there are an increasing number of Indonesian immigrants in America. The majority of them settled and immigrated to the United States after finishing their college education. Indonesia is the fourth most populous country in the world and home to the largest Moslem population in the world (The World Factbook, 2009). Therefore, Bahasa Indonesia and its related dialect, Bahasa Malau, is the most widely spoken language in Southeast Asia. Unfortunately, there is little phonetic research in the Indonesian language in general and even less in term of the acquisition of English by Indonesian children who were born in America. Many of these children grew up in a home environment in which both parents are Indonesian native speakers and late English learners. Therefore, these parents have distinct accent in their English pronunciation. One distinct pronunciation that can be seen is the pronunciation of stop consonants.

This study seeks to examine the phonetic characteristics of Indonesian children’s stop consonants pronunciation as compared to their Indonesian parents and American English monolingual children. Specifically these three questions will be addressed.
1. What are the mean VOT values of Indonesian children and adults who participated in this study.

2. How and to what extent the VOT characteristics of English stop consonants production differ between children who have Indonesian parents and monolingual native American children.

3. What are the possible contributing or determinant factors that make their pronunciation differ.

Unfortunately, there is no previous study that examines the same questions. Therefore, this study will serve as a starting point in attempt to understand the pronunciation of word-initial English stop consonants by Indonesian children primarily by analyzing voice onset time (VOT) of (/p/, /t/, /k/, /b/, /d/, /g/) pronounced by Indonesian children ages 2-9 years old. The result of mean VOT values for those stop consonants in English will be compared to mean VOT values of the same stop consonants in Indonesian language in a later study.
Chapter 2: Literature Review

There are some basic differences in Bahasa Indonesia and English stop pronunciation. In Bahasa Indonesia, /p/, /t/, and /k/ are pronounced without a release of air (aspiration). On the contrary, in the “standard” American English those consonants are pronounced with aspiration. Therefore, it is predictable that aspiration in these voiceless English stops will present problems for Indonesian speakers who are learning English. Most likely, Indonesian speakers will say /p, t, k/ in English also without the expected aspiration.

There are also place of articulation differences for stops between the two languages. The stops /t/ and /d/ are dental sounds made with constriction between the tip of the tongue and the upper teeth in Bahasa Indonesia. However, English pronunciation of /t/ and /d/ is alveolar meaning the blade of the tongue is on the alveolar ridge (just behind the upper teeth) (Soderberg, 2008).

Stop consonants are the most common consonants in both pre-linguistic and early linguistic development regardless of a child’s ambient language community (Kim & Gammon, 2010). In addition, more than 88% of languages use voicing to differentiate homorganic stops by exploiting three major timing relationships: lead voicing, long lag and short lag (MacLeod & Gammon, 2005). For the last five decades, VOT (voice onset time) has been used by researchers to measure the timing relationships that makes the distinctions among these three categories.

VOT

The important role of VOT to categorize stop consonants in a language was first proposed by Lisker and Abramson in 1964. VOT is defined as the temporal relation between the onset of glottal pulsing and the release of the initial stop consonant (Lisker & Abramson, 1964).
It is the length of time between the released of stop consonant to the start of voicing. VOT value can be negative when voicing start prior to consonant release. In order to measure the VOT value, the point of stop released and voicing must be determined.

In their study, Lisker and Abramson examined stop consonants data from 11 languages. They categorized stop consonants into 3 categories: lead, short lag and long lag. Those categories are differentiated by VOT values. Long lead is a category in which voicing start about 75-125 ms before the release of consonant (negative value VOT). Short lag is when voicing start 0-25 ms after the release. Long lag is when voicing start 60-100 ms after the release. (Lisker & Abramson, 1964).

Previous research has established that late learners are apt to produce English /p, t, k/ with VOT values that are too short for English (Flege, 1991). It would be interesting to see whether this condition would be true for Indonesian adult speakers in this study. In addition, it would be interesting to observe the possible differences that may occur.

**Voicing**

Another dimension that linguists use to distinguish stop consonant is “voicing.” Voicing is defined as the vibration of vocal cords. Therefore there is a presence of vibration when voiced consonants are being produced and there is an absence of vibration when voiceless consonants are being produced.

Bahasa Indonesia has similar stop consonant voicing contrast to English. Both languages have voiced and voiceless stop consonants. However there are two phonetic differences between stops in English and in Bahasa Indonesia. First of all, voiced stops in English can occur in all environments such as root-initial, root-medial, and root-final position. Whereas voiced stops in
Bahasa Indonesia can only occur in root-initial and root-medial positions. Second, voiceless stops in Bahasa Indonesia \([p,t,k]\) are unaspirated in root-initial position. However, these stops are pronounced with aspiration in English. In English, phoneticians argue that the voiced/voiceless distinction in initial position is actually an aspirated/unaspirated distinction.

(1) Example of stops in Bahasa Indonesia.

<table>
<thead>
<tr>
<th></th>
<th>Voiced stops</th>
<th>Voiceless stops</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a. In root-initial position</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[babi]</td>
<td>‘pig’</td>
<td>[pipi] ‘cheek’</td>
</tr>
<tr>
<td>[dada]</td>
<td>‘chest’</td>
<td>[tahu] ‘tofu’</td>
</tr>
<tr>
<td>[gajah]</td>
<td>‘elephant’</td>
<td>[kamar] ‘room’</td>
</tr>
<tr>
<td><strong>b. In root-medial position</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[sabun]</td>
<td>‘soap’</td>
<td>[papan] ‘wooden platform’</td>
</tr>
<tr>
<td>[bidan]</td>
<td>‘midwife’</td>
<td>[hitam] ‘black’</td>
</tr>
<tr>
<td>[sagu]</td>
<td>‘sagoo’</td>
<td>[luka] ‘laceration’</td>
</tr>
<tr>
<td><strong>c. In root-final position</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do not exist.</td>
<td></td>
<td>[santap] ‘formal word for eat’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[sikat] ‘brush’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[rasuk] ‘possessed by demon’</td>
</tr>
</tbody>
</table>
Chapter 3: Methodology

Subjects

Thirteen Indonesian adults and seven American adults were recruited for this study. Seven children whose both parents are monolingual American English speakers and thirteen children whose both parents are Indonesian-English bilingual speakers. In addition, one child has an American monolingual mother and an Indonesian-English bilingual father. Data from this particular subject is not included in the analysis of mean VOT and mean closure duration. Therefore, there were 40 participants in this study. All participants live in Columbus, Ohio.

Table 1. Subjects’ gender, birth country and home language preferences.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Gender</th>
<th>Birth Country</th>
<th>Home Language Preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult English Speaker</td>
<td>5 females, 2 males</td>
<td>America</td>
<td>7 English</td>
</tr>
<tr>
<td>Child English Speaker</td>
<td>4 females, 3 males</td>
<td>America</td>
<td>7 English</td>
</tr>
<tr>
<td>Adult Indonesian Speaker</td>
<td>9 female, 4 males</td>
<td>Indonesia</td>
<td>5 English, 6 Indonesia, 2 half</td>
</tr>
<tr>
<td>Child Indonesian</td>
<td>5 females, 7 males</td>
<td>America, 1 born in Indonesia.</td>
<td>5 English, 4 Indonesia, 4 half</td>
</tr>
<tr>
<td>Child with American mother &amp; Indonesian father</td>
<td>1 male</td>
<td>America</td>
<td>1 English</td>
</tr>
</tbody>
</table>

Indonesian children who were recorded are exposed to Bahasa Indonesia and Indonesian-accented-English at home through their bilingual parents. Parents of these children are native speakers of Bahasa Indonesia. The children’s age ranged from two and a half to nine years old. American children who were recorded were exposed only to native-English-speakers at home. These children should not be exposed to other languages intensively by any other caregivers. The children’s age ranged from three to eight years old. All of these children are typical developing children. The parents reported that none of the children have language delay, hearing problem, cognitive delay, or speech problem.
Table 2. Children’s ages

<table>
<thead>
<tr>
<th>Ages</th>
<th>Indonesian children</th>
<th>American children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages 2.5 years old</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ages 3 years old</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age 4 years old</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Ages 5 years old</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Ages 6 years old</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ages 7 years old</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Ages 8 years old</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>ages 9 years old</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

All of the Indonesian adult participants came to America for college education (age 17-20), except for one adult male Indonesian who came to America when he was 7 years old. All of the American English speakers are educated middle class adults living in Columbus Ohio.

In addition to recording utterances containing stop consonant, Indonesian parent-child pairs were also recorded producing spontaneous speech. They were told to speak and interact in Bahasa Indonesia for 5 minutes. Each conversation was recorded directly into a computer’s hard drive using Adobe Audition. This conversation was intended for future study and was not analyzed for this study.

All subjects were recruited personally by the principal investigator through her community involvement as well as emails and flyers. Any parent who identified his/her child as fitting the criteria was eligible to participate. Subjects were compensated $10 each for their participation.

**Preliminary procedure.**

Participants were escorted to Speech Perception and Acoustics Laboratories at The Ohio State University Speech and Hearing Science Department on the day of the recording. As soon as they were situated in the lab, the parent/caregiver was asked to complete a consent form and a permission form to participate in the study. The accompanied child was able to play or read book
in another room adjacent to the recording room while waiting for their parents. Upon completion of the consent forms, two background questionnaires (Appendix A for parent and Appendix B for child) were given to the parent to complete. It was necessary that the accompanying parent filled these questionnaires as accurately and with as much detail as possible.

The purpose of these forms was to enable the experimenter to obtain detail information about the child’s background such as country of birth, age, school, as well as linguistic background such as home language preference. The information were useful in differentiating Indonesian children who have been exposed to more English-speaking environments such as school to those that have minimum exposure such as younger children who remain in the care of their Indonesian mother. The information was also useful in differentiating those who speak mainly English at home, mainly Indonesian, or mixed half and half between Indonesian and English.

**Stimuli**

The recording material consists of two sets of words (Appendix D and Appendix E). One set contains Indonesian words which cover all the English stop consonants (/p/, /t/, /k/, /b/, /d/, /g/) in initial position (Appendix D). Another set consists of English words that have similar stop consonants and vowel combination (Appendix E). The five vowels used in the stimuli were /a/, /e/, /i/, /o/, and /u/. In addition there were words such as “ladder, water, muddy, city” versus “putar, kadal, kotor, kuda” and “spesial, stasiun and skolah” versus “sky, spot, and stair”. Therefore there were 37 words in each category. These additional words were not analyzed in this study. They were intended for future studies examining dental flaps in medial position in the two languages.
All of the stimuli were presented in an intervocalic environment. The English stop consonant stimuli were presented in a sentence preceded by “That’s a” (i.e. “That’s a book”). The Indonesian stop consonant stimuli were presented in a sentence preceded by “itu” (i.e. “Itu buku”). In intervocalic positions, speakers sustain the vibration of the vocal folds longer so that the low-frequency periodic energy lasts for an extended period of time of stop closure duration (Jacewicz, Fox, Lyle, 2009).

A picture of each stimuli appeared on the screen and participants named the pictures by reading the sentence below the picture. The picture was easy enough to identify even though the child was not able to read the sentence. In the case of young children who were unable to read, the experimenter gave explicit instructions and few examples that demonstrated “that’s a” must precede the word that the picture represented.

![That’s a bee](image)

*Figure 1. Sample of an English stimulus.*
This figure illustrates one of the English stimuli in the study.

**Data collection / Recording**

Participants were recorded in a sound attenuating booth in the Speech Perception and Acoustics Laboratories in the Department of Speech and Hearing Science at the Ohio State University. The participants were seated facing a computer screen inside the booth. A microphone on a stand was located on the table between the computer and the participant. Some
of the younger children sat in front of the computer with their caregiver and the experimenter. Participants were allowed to take as many breaks as they liked. Participants were asked to plan to be in the lab for two hours for the recording sessions.

During the recording session, the experimenter would click on the record button on the computer screen after the participant said each sentence. If a child was unable to name the word presented on the screen, the parent would say the sentence first. The children were told to repeat the sentence after the parent finished. In this case of prompted response, the response was recorded as “prompt response.” There was a special button image on the computer that the experimenter would be able to click to record the prompted response.

The participants were asked to record six practice words (Appendix C) in each category for a total of twelve words before recording the words from each language. The first language that the Indonesian speakers were asked to do is Bahasa Indonesia. Each subject was asked to go through each word list twice. There were thirty seven pictures in the list of Bahasa Indonesia stimuli (Appendix D). The same procedure was executed with the English picture set (Appendix E) after they were done with the Indonesian picture set. American English monolingual speakers were asked to record only the English picture set twice. Recording was done using a special Matlab program that displayed the pictures and recorded the utterances directly to a hard drive.

Data analysis

The recordings were analyzed using spectrographic analysis (utilizing Adobe Audition and TF32 software program). The starting point of stop closure and stop release during each stop were carefully measured and recorded. The duration between the starting point of stop closure and the released point were calculated as closure duration. VOT was the duration measured from
the point of closure released to the start of voicing. This acoustic measurement is illustrated in figure 2.

![Diagram of Stop Closure, Voicing Onset, Closure Duration, VOT, and Closure Release markers for calculating VOT and closure duration (the word: baby).](image)

*Figure 2.* Markers for calculating VOT and closure duration (the word: baby).

Then, voicing is carefully examined and recorded. Voicing during the stop was classified into four categories in this study. Voicing code 0 indicates lack of voicing during the entire closure duration. Voicing code 1 is assigned when voicing started prior to closure release. Voicing code 2 is assigned when there was voicing throughout the entire closure duration. Voicing code 3 is assigned when there is partial voicing in the initial part of closure duration but stopped before closure released. The voicing code categories can be seen in figures 3-6.
The first category of voicing is called voiceless and not pre-voiced in which no voicing occurs during closure duration (voicing code 0) as shown in figure 3.

![Figure 3](image)

*Figure 3. An example of no voicing on a spectrogram (the word: tub).*

The second type of voicing is called pre-voicing (voicing code 1). Pre-voicing occurs when voicing is evident before the stop released. The pre-voicing category is shown in figure 4.

![Figure 4](image)

*Figure 4. An example of pre-voicing on a spectrogram (the word: doddle).*
The third type of voicing (voicing code 2) is a complete voicing in which voicing occurs throughout the entire closure duration, shown in figure 5.

![Complete voicing](image)

Figure 5. An example of complete voicing on a spectrogram (the word: baby).

The last type of voicing is partial voicing in which voicing occurs only in the initial part of closure duration (voicing code 3) often caused by longer closure duration as shown in figure 6.

![Partial voicing](image)

Figure 6. An example of partial voicing on a spectrogram (the word: duck).
The locations of stop closure, closure release, the start of voicing and voicing code were measured and the measurements were entered into an Excel file. These recorded time allow us calculate a closure duration and VOT. All measurements were converted to ms.

Closure duration: (closure released time – stop closure time)*1000 ms

VOT: (Time of voicing after closure released-closure released)*1000ms.

Finally, all the data were analyzed using SPSS software.
Chapter 4: Result

Voicing

As cited by Jacewicz, Fox & Lyle, 2009, there is evidence from developmental studies that young children’s voiced stops are often devoiced and have less closure voicing than adults (Kewley-Port & Preston 1974). This is evident in the result shown in table 3. In addition, a stop occurring in the intervocalic position should naturally be voiced throughout most if not all of the closure if it is short. In the case of a longer closure the stop will be voiced though the initial portion of the closure and then become voiceless (Jacewicz, Fox & Lyle, 2009). Indonesian adult and children participants have longer closure duration and therefore they have higher number of stops that are initially voiced.

Table 3. Result of voicing categories across speakers.

<table>
<thead>
<tr>
<th>Stop Voicing</th>
<th>Speakers</th>
<th>Voicing Code</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>VOICELESS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult English</td>
<td>70 (67%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Child English</td>
<td>67 (56%)</td>
<td>1 (.85%)</td>
<td>2 (1.7%)</td>
</tr>
<tr>
<td>Adult Indonesian</td>
<td>155 (79.5%)</td>
<td>1 (.5%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Child Indonesian</td>
<td>140 (80%)</td>
<td>2 (1%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>432 (73%)</td>
<td>4 (.67%)</td>
<td>2 (.33%)</td>
</tr>
</tbody>
</table>

| VOICED          |           |   |   |   |   |
| Adult English | 4 (4%)  | 0 (0%) | 75 (71%) | 26 (25%) | 105 (100%) |
| Child English | 22 (19%) | 0 (0%) | 35 (30%) | 61 (51%) | 118 (100%) |
| Adult Indonesian | 28 (14.50%) | 1 (.51%) | 101 (52%) | 63 (33%) | 193 (100%) |
| Child Indonesian | 53 (30%) | 2 (1%) | 34 (24%) | 79 (45%) | 176 (100%) |
| TOTAL            | 107 (17.5%) | 3 (.5%) | 253 (43%) | 229 (39%) | 592 (100%) |
It can be seen in the table that adult Indonesian speakers and child Indonesian speakers have more voiced stops that are partially voiced in the beginning of the closure duration. Adult Indonesian speakers have 63% partially voiced consonants and child Indonesian speakers have 79% partially voiced consonants. This was caused by the fact that adult Indonesian speakers and child Indonesian speaker have longer closure duration. In addition child English speakers also have higher percentage of partially voiced consonants of 61% compare to adult English speakers who only have 26%. Child English speakers have considerably longer closure duration than adult English speakers.

**Mean VOT and closure duration across speaker types.**

There are no significant differences across VOT values for all speaker types in their voiced and voiceless stop consonants pronunciation as shown in table 4 and represented in the graph in figure 7. However, both Indonesian adult and child speakers have longer closure duration compared to English monolingual speakers as shown in table 5 and represented in figure 8.

Table 4. Mean VOT across speaker types.

<table>
<thead>
<tr>
<th>Spk_type</th>
<th>stop_voicing</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult English</td>
<td>Voiceless</td>
<td>53.8</td>
<td>105</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>17.4</td>
<td>30</td>
<td>9.6</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>45.7</td>
<td>135</td>
<td>21.6</td>
</tr>
<tr>
<td>Child English</td>
<td>Voiceless</td>
<td>72.8</td>
<td>101</td>
<td>29.5</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>24.0</td>
<td>83</td>
<td>17.4</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>50.8</td>
<td>184</td>
<td>34.7</td>
</tr>
<tr>
<td>Adult Indonesian</td>
<td>Voiceless</td>
<td>43.3</td>
<td>195</td>
<td>19.2</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>20.9</td>
<td>92</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>36</td>
<td>287</td>
<td>19.8</td>
</tr>
<tr>
<td>Child Indonesian</td>
<td>Voiceless</td>
<td>71.5</td>
<td>175</td>
<td>35.9</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>24</td>
<td>134</td>
<td>18.9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>51</td>
<td>309</td>
<td>37.9</td>
</tr>
<tr>
<td>Total</td>
<td>Voiceless</td>
<td>59</td>
<td>576</td>
<td>29.8</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>22.6</td>
<td>339</td>
<td>15.9</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>45.5</td>
<td>915</td>
<td>31</td>
</tr>
</tbody>
</table>
This result suggested that the L1 English learners (Indonesian children who were born in America) developed a phonetic system for stop consonants that are similar to American English monolingual children. On the contrary, Indonesian adults who were born in Indonesia and were exposed to English consistently as adults (L2 English learners) have different phonetic characteristics.

*Figure 7.* Mean VOT across speaker types. This figure compares mean VOT values for voiced and voiceless stops across speaker types.

In the meantime, closure duration is considerably longer across Indonesian adult and child speakers as compared to English adult and child speakers as shown in table 5.
Table 5. Closure duration across speaker types for voiced and voiceless stops.

<table>
<thead>
<tr>
<th>Spk_type</th>
<th>stop_voicing</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult English</td>
<td>Voiceless</td>
<td>70.4</td>
<td>105</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>72.5</td>
<td>105</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>71.4</td>
<td>210</td>
<td>16.6</td>
</tr>
<tr>
<td>Child English</td>
<td>Voiceless</td>
<td>99.4</td>
<td>104</td>
<td>41.0</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>111.3</td>
<td>103</td>
<td>37.6</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>105.3</td>
<td>207</td>
<td>39.7</td>
</tr>
<tr>
<td>Adult Indonesian</td>
<td>Voiceless</td>
<td>105.3</td>
<td>195</td>
<td>29.2</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>99.6</td>
<td>193</td>
<td>26.6</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>102.5</td>
<td>388</td>
<td>28.0</td>
</tr>
<tr>
<td>Child Indonesian</td>
<td>Voiceless</td>
<td>126.8</td>
<td>175</td>
<td>46.3</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>131.8</td>
<td>176</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>129.3</td>
<td>351</td>
<td>55.3</td>
</tr>
<tr>
<td>Total</td>
<td>Voiceless</td>
<td>104.4</td>
<td>579</td>
<td>40.5</td>
</tr>
<tr>
<td></td>
<td>Voiced</td>
<td>106.6</td>
<td>577</td>
<td>46.6</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>105.5</td>
<td>1156</td>
<td>43.7</td>
</tr>
</tbody>
</table>

Figure 8. Mean closure duration across speaker types. This figure compares mean closure duration values for voiced and voiceless stops across speaker types.
Mean VOT for Indonesian adults based on home language preference.

The data shown previously showed that mean VOT for English adult speakers is 17.4 ms for voiced stop consonants and 53.8 ms for voiceless stops. Table 6 shows voiced and voiceless mean VOT for Indonesian adult speakers based on their home language preferences.

Table 6. Mean VOT based on home language preferences for Indonesian adults.

<table>
<thead>
<tr>
<th></th>
<th>Indonesian</th>
<th>Half and Half</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>20.8</td>
<td>24.2</td>
<td>19.2</td>
</tr>
<tr>
<td>Voiceless</td>
<td>43.3</td>
<td>40.8</td>
<td>44.5</td>
</tr>
</tbody>
</table>

It is interesting to see that Indonesian adults who speak English at home have closer mean VOT value to English monolingual adult speakers compare those who speak Bahasa Indonesia at home. Moreover, Indonesian adult speakers who speak half English and half Bahasa Indonesia at home have the biggest value differences to the mean VOT value of English monolingual adult speakers. This is an interesting fact to explore in further detail because the higher value differences could be caused by language interference between L1 and L2.

Mean closure duration for Indonesian adults based on home language preferences.

In addition, English adult speakers’ mean closure duration for voiced consonant was 72.5 ms and for voiceless consonant was 70.4 ms. Mean closure duration for Indonesian adult speakers were considerably longer than those of English monolingual adult speakers as shown in table 7.

Table 7. Mean closure duration based on home language preferences for Indonesian adults.

<table>
<thead>
<tr>
<th></th>
<th>Indonesian</th>
<th>Half and Half</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>99.5</td>
<td>90.5</td>
<td>104.3</td>
</tr>
<tr>
<td>Voiceless</td>
<td>99.6</td>
<td>113.2</td>
<td>111.3</td>
</tr>
</tbody>
</table>
Mean VOT for Indonesian children based on home language preference.

Mean VOT value for English monolingual children was 24 ms for voiced consonant and 72.8 ms for voiceless consonant. Table 8 shows mean VOT value for Indonesian children based on their home language preferences.

Table 8. Mean VOT based on home language preferences for voiced and voiceless stop consonant for Indonesian children.

<table>
<thead>
<tr>
<th></th>
<th>Indonesian</th>
<th>Half and Half</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>32.9</td>
<td>17.4</td>
<td>25.7</td>
</tr>
<tr>
<td>Voiceless</td>
<td>50</td>
<td>73.7</td>
<td>93.9</td>
</tr>
</tbody>
</table>

It is interesting to see that VOT values for children who prefer to speak English at home is similar to VOT value of English monolinguals children’s in voiced consonants. On the other hand, it is surprising to see that VOT values for voiceless consonants produced by Indonesian children who speak half English and half Indonesian at home is similar to VOT value of monolingual child English speakers. It was expected that those who prefer to speak Indonesian at home will have similar VOT values with Indonesian adult speakers who also prefer to speak Indonesian at home.

Mean closure duration for Indonesian children based on home language preferences.

Mean closure duration for Indonesian children are considerably higher across Indonesian and English home language preferences as compared to English children speakers. Nevertheless, the mean closure duration of Indonesian children who speak half and half is similar closure duration value with English monolingual speakers and Indonesian adults who also speak half and half.
Table 9. Mean closure duration based on home language preferences for voiced and voiceless stop consonants for Indonesian children.

<table>
<thead>
<tr>
<th></th>
<th>Indonesian</th>
<th>Half and Half</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiced</td>
<td>144.3</td>
<td>112.6</td>
<td>148.6</td>
</tr>
<tr>
<td>Voiceless</td>
<td>148.4</td>
<td>118.5</td>
<td>114.2</td>
</tr>
</tbody>
</table>
Chapter 5: Discussion.

The Indonesian adult speakers might have different phonetic characteristics in their production of stops because they had passed a critical period for learning the language that would make them fluent as native speakers. Lenneberg, said “that a foreign accent in an L2 is “inevitable” if it is learned after puberty because brain development and lateralization for language function have reached completion by that time. (Lenneberg, 1967).

It was mentioned earlier in the introduction that Indonesian voiceless stop consonants in initial position are not aspirated, contrary to voiceless aspirated English stop consonants in initial position. Therefore, it was predicted that Indonesian adult speakers who are late English learners will have difficulty pronouncing the aspirated stops in initial position. This prediction is supported by the VOT findings in voiceless stop consonants as shown in figure 9.

![Figure 9. Mean VOT for each consonant across speaker types. This figure compares mean VOT values for each consonant across speaker types.](image-url)
Figure 9 shows that (/p/,/t/,/k/) in the Indonesian adult speakers has shorter VOT compared to the American adult speakers. However, Indonesian children VOT in (/p/ and/k/) are not much different than the American monolingual English children. In fact, the mean VOT for /t/ is higher in Indonesian children than in English children for reason yet to be investigated.

The differences between early English learners (Indonesian children born in America) and late English learners (Indonesian adult parents) VOT values can also be explained using the speech learning model (SLM). The Speech Learning Model was first described by Flege in 1988. The SLM posits that speech production is organized at phonemic category, phonetic category, and sensory-motor levels (Flege, 1991). Therefore, after the phonetic categories have been established for L1 sounds in early childhood, listeners are increasingly likely to identify L2 sounds that are partially resemble corresponding sounds in the L1 as being realizations of an L1 category (Flege, 1991).

Therefore in this case, Indonesian adults as late English learners will continue to pronounce similar sounds (stops in English and Bahasa Indonesia) with minimal distinctions. Nevertheless, Indonesian children who were born in American and had been exposed to English earlier in life will eventually able to differentiate the acoustic phonetic differences for stops in English and in Bahasa Indonesia.

With regard to home language preferences, very few studies have specifically examined this condition. One of the studies was conducted by Goldstein, Fabiano and Washington in 2005. The result of their study indicated that there were no significant correlations between amount of output in each language and phonological skill (Goldstein, Fabiano & Washington, 2005). The result of this study showed slight differences in mean values between
speakers based on home language preferences. However, the values may not be significant enough to support a hypothesis that home language preferences can be a contributing factor in determining the phonetic contrast between English and Bahasa Indonesia spoken by Indonesian children.

Limitation in this study is that there is no known data available regarding Indonesian adult or children phonetic characteristic of stops that this study can compare its result with.
Chapter 6: Summary/Conclusion

In general, documenting bilingual children’s phonetic development is a very difficult and challenging task. It is important to remember that the result of this study might be influenced by important factors such as rate of phonetic context, age of children, and stress position. In addition, there is very little data available on Indonesian children phonetic acquisition of English in particular and of Bahasa Indonesia in general to compare the result of this study.

Therefore, this study serves as a pilot study on American-born-Indonesian-children’s language acquisition research. At this point, it would be too early to draw a solid conclusion on how Indonesian children acquire English phonetic system. Therefore, further study is crucial to learn more extensively on how these two languages interact in Indonesian children who were born in America.

Nevertheless, the result of this study support the critical period hypothesis that said adult learners of an L2 will not be able to develop “native” like ability in pronunciation compare to early learners of the L2. The VOT differences between Indonesian adult speakers and Indonesian children speakers reflected the evidence for the critical period hypothesis because Indonesian children speakers have similar VOT values as their English children monolingual speakers.

As of the case of voicing and closure duration, further research need to be done to compare the closure duration value. Therefore, it is imperative that this study should be follow by another study that will examine Indonesian monolingual adult speakers and Indonesian monolingual children speakers. These future research participants should not be exposed to English language extensively for more than 6 months.
References


Appendix A: Parent Background Questionnaire.

**LANGUAGE BACKGROUND QUESTIONNAIRE (ADULT)**

1. Date of birth ________________________  
   Current age __________

2. Sex  
   MALE  
   FEMALE

3. Where have you lived, and what age were you at the time (starting with place of birth)?  
   *Place (city, area, country)*  
   *Ages (from ___ to ___)*

<table>
<thead>
<tr>
<th>Place (city, area, country)</th>
<th>Ages (from ___ to ___)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Where were your parents or other caretakers born? Where did they grow up? What languages do/did they speak?  

<table>
<thead>
<tr>
<th>Relation</th>
<th>Place of birth</th>
<th>Place(s) they grew up</th>
<th>Language(s) spoken</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. What language(s) did your parents or other caretakers speak to you at home when you were growing up?  

   __________________________________________________________

6. What do you consider to be your native language(s)?  
   ______________________________
7. Including your native language(s), what languages do you know? At what age did you begin learning each language? How well can you write, read, speak, and understand each language?

<table>
<thead>
<tr>
<th>Language</th>
<th>Age</th>
<th>Writing</th>
<th>Reading</th>
<th>Speaking</th>
<th>Understanding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
<td>1 2 3 4</td>
</tr>
</tbody>
</table>

8. For each language you mentioned in the question above, please estimate the percentage of your current language use that takes place in each language.

<table>
<thead>
<tr>
<th>Language</th>
<th>% of current language use</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. Do you interact with speakers of English from other parts of the world (India, Great Britain, Indonesia, Singapore, etc.)? YES NO

If you circled yes, please describe who, how often, and where each speaker is from.

________________________________________________________________________
________________________________________________________________________

10. Do you interact with speakers of foreign-accented English? YES NO

If you circled yes, please describe who, how often, and each speaker’s native language(s).
11. Circle the highest level of education you have completed so far.

   PRIMARY   JUNIOR HIGH   HIGH SCHOOL   COLLEGE   POST-GRADUATE

12. What is your profession? (If academic or student, please indicate field of study.)

__________________________________________________________________________

13. Do you have any speech, language, or hearing disorders?  YES  NO
   If you circled yes, please provide details.

__________________________________________________________________________

__________________________________________________________________________
Appendix B: Child Background Questionnaire

LANGUAGE BACKGROUND QUESTIONNAIRE (CHILD)

1. Child’s date of birth ________________________               Child’s current age __________

2. Child’s sex               MALE           FEMALE

3. Where has your child lived, and what age was s/he at the time (starting with place of birth)?
   Place (city, area, country)               Ages (from ___ to ___)
   ______________________________________  ______________________________________
   ______________________________________  ______________________________________
   ______________________________________  ______________________________________
   ______________________________________  ______________________________________
   ______________________________________  ______________________________________

4. What language(s) does your child hear and spoken at home?
   _____________________________________________________________

5. If your child hears more than one language at home, please estimate the percentage of time s/he
   hears each language, relative to the other language(s).

   Language       % of current language experience
   _______________       __________________
   _______________       __________________
   _______________       __________________
   _______________       __________________
   _______________       __________________

6. What do you consider to be your child’s native language(s)?    __________________________
7. Does your child interact with speakers of English from other parts of the world (India, Great Britain, Indonesia, Singapore, etc.)?  
   YES  NO
   If you circled yes, please describe who, how often, and where each speaker is from.

________________________________________________________________________

________________________________________________________________________

8. Does your child interact with speakers of foreign-accented English?  
   YES  NO
   If you circled yes, please describe who, how often, and each speaker’s native language(s).

________________________________________________________________________

________________________________________________________________________

9. Does your child attend daycare/preschool/kindergarten?  
   YES  NO
   If you circled yes, please indicate how many hours per week your child attends each of the following:
   
   In-home care (relative, friend, neighbor, etc.) ________
   Commercial daycare center/preschool program ________
   Elementary school kindergarten class ________

   If you circled yes, please indicate what language your child’s daycare providers or teachers speak when interacting with your child.

________________________________________________________________________

10. Does your child have any speech, language, or hearing disorders?  
    YES  NO
    If you circled yes, please provide details.
Appendix C: Practice Stimuli

- That’s a baseball
- That’s a cape
- That’s a car
- That’s a dog
- That’s a key
- That’s a pool
- Itu bakmi
- Itu gelas
- Itu topi
Itu piring

Itu tas

Itu kompor
Appendix D: Bahasa Indonesia stimuli

- Itu babi
- Itu bebek
- Itu bibir
- Itu bola
- Itu bulan
- Itu dada
- Itu dede
- Itu dingin
- Itu dompet
Stop Consonant Characteristics

- Itu duduk
- Itu gajah
- Itu gelap
- Itu gigi
- Itu geli
- Itu gala
- Itu kadai
- Itu kamar
- Itu kentut
Stop Consonant Characteristics

1. Itu tomat
2. Itu tidur
3. Itu tutup pintu
4. Itu teko
5. Itu kotor
6. Itu putar-putar
7. Itu kodok
8. Itu stasiun
9. Itu sekolah
Appendix E: English Stimuli

That's a baby

That's a book

That's a bee

That's a boat

That's a bottle

That's a cap

That's a cat

That's a coat

That's a cookie
That's a kitchen

That's a deer

That's a desk

That's a donut

That's a dodeliebops

That's a duck

That's a geese

That's a girl

That's a goat
That's a goose

That's a gun

That's a pencil

That's a pig

That's a pony

That's a poodle

That's a puddle

That's a table

That's a tea cup
That's a stair