Investigating the Prevalence of Use by Japanese Speakers of an Acceptable Alternative Articulation of the Phoneme /s/ to That Commonly Taught in ESL and EFL Classrooms

Greg Raver-Lampman

Follow this and additional works at: https://digitalcommons.odu.edu/english_etds

Part of the Applied Linguistics Commons, English Language and Literature Commons, First and Second Language Acquisition Commons, and the Phonetics and Phonology Commons
INVESTIGATING THE PREVALENCE OF USE BY JAPANESE SPEAKERS OF AN ACCEPTABLE ALTERNATIVE ARTICULATION OF THE PHONEME /S/ TO THAT COMMONLY TAUGHT IN ESL AND EFL CLASSROOMS

by

Greg Raver-Lampman
B.A. May 1979, University of California, Berkeley

A Thesis Submitted to the Faculty of Old Dominion University in Partial Fulfillment of the Requirements for the Degree of

MASTER OF ARTS

APPLIED LINGUISTICS

OLD DOMINION UNIVERSITY
May 2012

Approved by:

Janet Bing (Director)

Joanne Scheibman (Member)

Felicia Toreno (Member)
ABSTRACT

INVESTIGATING THE PREVALENCE OF USE BY JAPANESE SPEAKERS OF AN ACCEPTABLE ALTERNATIVE ARTICULATION OF THE PHONEME /s/ TO THAT COMMONLY TAUGHT IN ESL AND EFL CLASSROOMS

Greg Raver-Lampman
Old Dominion University, 2012
Director: Dr. Janet Bing

The International Phonetic Association (IPA) as well as textbooks on phonology and teaching English as a second language (ESL) or foreign language (EFL) characterize the /s/ as an “alveolar fricative,” meaning that the tongue approaches the alveolar ridge to produce the sound. Japanese phonology texts characterize the Japanese /s/ as alveolar as well. This tongue position has become integral to teaching the sound to English-speaking children who have speech impediments and for teaching the sibilants to speakers of other languages, including first-language speakers of Japanese who often struggle with the English /s/ despite the fact that the sound occurs in their own language. A recent study suggests that many first-language English speakers use an acceptable alternative mandibular articulation of the /s/, with the apex of the tongue touching their lower incisors. At article by an EFL teacher in Japan posited that Japanese speakers may pronounce the /s/ with this acceptable alternative in their own language. To date, no research has examined the articulation of the Japanese /s/ with technology that can show whether the Japanese do, in fact, pronounce the /s/ with their tongues in the alveolar position or with the acceptable alternative placement used by to pronounce the English /s/.

This study used ultrasound to produce visual images of the tongue movement of 20 Japanese first-language speakers pronouncing words containing /s/ in their own
language in word-initial and word-medial positions. Results determined that 85 percent of the participants used an alternative articulation with the tongue at the lower incisors. Implications for instruction are discussed.
I dedicate this thesis to my wife, Sharon, and my daughter, Emmy, with whom I have shared a life of cultural and linguistic adventure.
ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and encouragement of Drs. Janet Bing and Joanne Scheibman, who encouraged me to research my own alternative articulation after a phonology course left me fearing I pronounced the /s/ sound incorrectly. Both have encouraged me to ask questions and pursue research where answers seemed unsatisfactory. This occurred not only with the subject of this thesis, but with every linguistic anomaly I stumbled across. Thanks to them, I believe I have increasingly sought out areas where the answers appear unsettled, with an eye toward fleshing them out.

I owe an enormous debt of gratitude to Dr. Felicia Toreno, Director of Diagnostic Medical Sonography at Tidewater Community College. When I decided to research the tongue position used by Japanese pronouncing the /s/ sound in their own language, ultrasound technology appeared to be the only feasible method. Linguistic ultrasonic research is an emerging field and has only been performed at well-funded centers where researchers could afford equipment that runs into tens of thousands of dollars. I sought sources for ultrasound devices and would have been forced to abandon this research if Dr. Toreno had not agreed to help. Over months of correspondence, Dr. Toreno and I worked out a considerable number of technical details and reviewed other ultrasonic linguistic research papers. Once we decided on a method to use ultrasound in this research, Dr. Toreno spent hours over three weekends at the research site in Newport News to perform oral ultrasound on 47 participants. She also gave me insight on how to examine the images produced. Her help, guidance and expertise made this thesis possible.
I also want to acknowledge the assistance of Atsuko Smith, a translator who translated the study materials into Japanese and also helped identify a venue where we could find enough Japanese-speaking participants to conduct this research. This task involved dozens of emails, telephone calls and one-on-one meetings, as well as a few dead ends. Mrs. Smith not only came to understand the research, she became invested in it. During the research, she provided live translation for the participants whose English was often extremely limited.

A final debt of gratitude goes to Mikihiko Sugaya, a teacher and principal of the Newport News Japanese School. Mrs. Smith and I met with him at his Chesapeake home over biscuits and tea and he wanted to know everything about the research. Once he was convinced of its value, Sensei Sugaya petitioned for approval from the board of the Newport News Japanese School, set aside a room for us, and recruited parents and teachers to participate at the rate of 15 each day.

In this project, the research would not have been possible without overcoming enormous logistical challenges surmounted with the help of these people who became vital parts of the research team.
TABLE OF CONTENTS

LIST OF TABLES ................................................................................................................ ix

LIST OF FIGURES .............................................................................................................. x

LIST OF GRAPHS ............................................................................................................. xiv

Chapter

1. INTRODUCTION AND BACKGROUND TO THE STUDY ...................................... 1
   1.1. INTRODUCTION AND REVIEW OF THE LITERATURE ...................... 1
   1.2. REVIEW OF MODALITIES TO CAPTURE DYNAMIC ARTICULATION ........................................... 7

2. METHOD AND PROCEDURE ................................................................................... 10
   2.1. USE OF ULTRASOUND ........................................................................... 10
   2.2. LIMITATIONS OF ULTRASOUND ......................................................... 11
   2.3. CREATION OF A POINT OF REFERENCE ............................................ 11
   2.4. ACOUSTIC SHADOW AND ATTENUATED REFLECTION ............... 12
   2.5. SYNCHRONIZATION OF VIDEO TO AUDIO TRACK ...................... 15
   2.6. ULTRASOUND DEVICE EMPLOYED ................................................... 16
   2.7. SONOGRAPHER ....................................................................................... 18
   2.8. STUDY MATERIALS ............................................................................... 19
   2.9. SETTING .................................................................................................... 21
   2.1.0. DATA SELECTION CRITERIA AND DEMOGRAPHIC CHARACTERISTICS OF PARTICIPANTS ............................................. 21
   2.1.1. STUDY PROCEDURE .............................................................................. 23
   2.1.2. PRESERVATION OF THE DATA ............................................................ 24

3. ANALYSIS OF THE DATA ........................................................................................ 26
   3.1. ANALYSIS OF THE ULTRASONIC IMAGES AND SOUND FILES ... 26
   3.2. INSUFFICIENT DATA ON /S/ BETWEEN ALVEOLAR AND MANDIBULAR VOWEL .......................................................... 37

4. RESULTS ..................................................................................................................... 39

5. DISCUSSION AND IMPLICATIONS FOR INSTRUCTION..................................... 45
   5.1. DISCUSSION AND IMPLICATIONS FOR INSTRUCTION .................. 45
   5.2. LIMITATIONS ........................................................................................... 47
   5.3. FUTURE RESEARCH AND CONCLUSION ........................................... 47
REFERENCES ................................................................................................................... 49

APPENDICES .................................................................................................................... 55
   A. STUDY WORD FOR ENGLISH SPEAKERS ................................................. 55 ..
   B. STUDY WORDS IN JAPANESE ................................................................. 56
   C. DEMOGRAPHIC SURVEY .............................................................................. 57
   D. DEMOGRAPHIC SURVEY IN JAPANESE .................................................... 58

VITA ................................................................................................................................... 59
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Study Words with Anglicized Spelling Followed by Phonetic Spelling and English Translation</td>
<td>20</td>
</tr>
<tr>
<td>2. Phonemic Environments of the /s/ in Utterances Used in the Study</td>
<td>20</td>
</tr>
<tr>
<td>3. Demographic Survey</td>
<td>22</td>
</tr>
<tr>
<td>4. Breakdown of /s/ Sounds in Word-initial and Word-medial Positions by Participant</td>
<td>40</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figures                                                                                       Page
1. Java-Enabled Interactive Illustration of IPA-Described Manner and Placement of the Sound /s/ from the University of Toronto ........................................... 2
2. Frame of X-ray Video Used in Phonetics Course (Ladefoged 2012)........................................ 9
3. Mindray C5 2S Transducer ........................................................................................................... 10
4. Annotated Illustration of an Ultrasound Image of the Oral Cavity Used in This Study with the Tongue in the Neutral Position, Neither Touching the Roof of the Mouth nor Depressed to the Floor of the Oral Cavity ........................................... 12
5. Participant 3 (Alveolar Producer), Frame 268........................................................................... 13
6. Participant 9, Frame 204............................................................................................................. 14
8. Participant 3, Frame 174. The Dorsum of the Tongue Rises Toward the Soft Palate to Form the [k] of Kasa ............................................................................................................. 17
9. Participant 3, Frame 176. The Tongue Descends Rapidly from the Soft Palate at the Release of the [k] in Kasa, Creating the /k/ Sound ............................................. 18
10. Participant 1, Frame 155. Pronouncing the Word-Final [o] with the Tongue on the Floor of the Oral Cavity .................................................................................................................. 27
11. Participant 9 (Mandibular Producer), Frame 1. Tongue in the Neutral Position............................................. 28
12. Participant 9, Frame 71. Tip of Tongue Descends to Form First [s] of Sasaki ................................................................. 29
13. Participant 9, Frame 79. Tongue Drops Further to Form the Second [s] of Sasaki, Formed after the Vowel [a] ................................................................. 29
14. Participant 9, Frame 90. Dorsum of the Tongue Vanishes as it Makes Contact with the Soft Palate to Articulate the [k] of Sasaki ................................................................. 30
15. Participant 9, Frame 98. Back of Tongue Drops, Apex Rises to Form the High Front Vowel [i] of Sasaki ................................................................. 30

16. Participant 9, Frame 112. Tongue Tip Descends from the [i] to Form the [s] of San ........................................................................................................ 31

17. Participant 9, Frame 128. Tongue Tip Moves Upward Through the Nasalized [ã] and Pushes Forward and Upward to Form the [n] of San ..................................................... 31

18. Participant 9, Frame 150. Tongue Tip Descends from the [n] to Form the [s] and Coarticulated [o] of Soko ........................................................................ 32

19. Participant 9, Frame 161. The Dorsum of the Tongue Rises and Vanishes as it Makes Contact with the Soft Palate to Form the [k] .................................................... 32

20. Participant 9, Frame 166. The Tongue’s Apex Pushes Forward and Upward to Form the [n] of Ni ........................................................................ 33

21. Participant 9, Frame 172. The Tongue Rises to Form the High Front Vowel [i] of Ni ......................................................................................... 33

22. Participant 9, Frame 185. The Tongue Tip Descends to Form the [s] of Suwatte ...................................................................................... 34

23. Participant 9, Frame 187. The Tongue Descends Further to Form the [wa] of Suwatte Toward the Floor of the Oral Cavity ........................................ 34

24. Participant 9, Frame 201. The Tongue Tip Rises to the Back of the Front Teeth to Form the Dentalized Articulation of the [t] of Suwatte ...................... 35

25. Participant 9, Frame 204. The Dorsum of the Tongue Vanishes as it Comes into Contact with the Soft Palate to Form the [k] of Kudasai ................. 35

26. Participant 9, Frame 212. The Tongue’s Apex Rises Toward the Top Teeth Forming the Lamino-Alveolar [d] of Kudasai ........................................................................ 36

27. Participant 9, Frame 220. The Tongue Descends to Form the [s] of Kudasai ........................................................................................................ 36


29. Participant 1, Frame 155. Mandibular Producer Pronouncing the [o] of Miso with the Placement Used by Participants When the Vowel Appeared in Word-Final Position 38

30. Participant 8 (Alveolar Producer), Frame 182 ........................................................................ 41
31. Participant 3 (Alveolar Producer), Frame 268 .............................................................. 41

32. Participant 3, (Alveolar Producer) Frame 283 .............................................................. 42

33. Participant 3 (Alveolar Producer), Frame 268. The Body and Apex of the 
Tongue Rise Toward the Palate to Produce the [s] of Kudasai. ..................................... 43

34. Participant 9 (Mandibular Producer), Frame 220. Tongue Descends Toward the 
Floor of the Oral Cavity to Form the [s] of Kudasai. ................................................. 43

35. Participant 3 (Alveolar Producer), Frame 283. The Tongue Lowers from the 
Alveolar Ridge to Eliminate the Turbulent Sound of the Fricative to Pronounce 
the High Front Final Diphthong of Kudasai ............................................................... 44

36. Participant 9 (Mandibular Producer), Frame 235. Tongue Ascends to Form 
the High Front Final Vowel of the Diphthong [ai] of Kudasai. The Place of 
Production of the Vocalic [ai] in the Mandibular Producer Differs from 
the Placement by the Alveolar Producer ................................................................. 44
<table>
<thead>
<tr>
<th>Graphs</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Results of Study by Participants Using IPA-Described Alveolar Articulation or the Acceptable Alternative Mandibular Articulation</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION AND BACKGROUND TO THE STUDY

1.1. INTRODUCTION AND REVIEW OF THE LITERATURE

The International Phonetic Association (IPA) as well as most textbooks on phonology and teaching English as a second language (ESL) or a foreign language (EFL) characterize the /s/ sound as a voiceless alveolar fricative, meaning that the tongue approaches the alveolar ridge to produce the sound (Celce-Murcia, 2001; Finegan, 2008; Hayes, 2009; Ladefoged, 1993). Japanese phonology texts describe the Japanese /s/ as alveolar as well (Vance, 1987). Studies using electropalatography, in which electrodes record the points at which the tongue comes in contact with the roof of the mouth, appeared to confirm the alveolar place of articulation of the English /s/ (Cheng, Murdoch, Goozee, & Scott, 2007). X-ray images of an English-speaking subject pronouncing a word with a word-medial /s/ also seemed to confirm the alveolar placement of the consonant (Ladefoged, 2012). Other tongue placements used in pronouncing the /s/ are often described as misarticulations or as defective articulations (Daniloff, 1980; McAuliffe & Cornwell, 2008; McGlone & Proffit, 1973; Mosher, 1929; Stephens & Daniloff, 1974).

The phonetic alphabet and chart describing the place and manner of production of sounds, including the /s/, date back to 1888, when linguist Paul Passy, founder of the IPA, published an alphabet that contained the sounds of spoken languages (MacMahon, 2010). By 1912, the IPA alphabet promulgated by Passy became the most widely used notation system and a standard part of both linguistic education and linguistic teaching.
(International Phonetic Association, 1912). While the IPA articulatory positions were meant to be descriptive illustrations of how these sounds are commonly produced, texts for EFL and ESL teachers often present these descriptions prescriptively, as the correct way to pronounce them (Becker, 2010; Celce-Murcia, 2001; Celce-Murcia, Brinton, & Goodwin, 2004; Finegan, 2008; Goodwin & In, 2001). Because most ESL and EFL teachers seldom progress beyond introductory phonology courses at the master’s level, suggestions that there is a correct articulatory position can leave them ill-equipped to teach acceptable alternative articulations that may be more effective for select populations (Cairns, 1988; Stoner, Gateley, & Rivers, 1987). The tendency to teach the IPA articulatory positions prescriptively has been exacerbated by technology, including web-based pronunciation guides (Becker, 2010) as well teaching aids including flash animations of the phonemes of the IPA chart developed at the University of Iowa (University of Iowa, 2010) and the Java-enabled interactive sagittal section (Hall, 2012) developed at the University of Toronto (see Figure 1).

Figure 1. Java-Enabled Interactive Illustration of IPA-Described Manner and Placement of the Sound /s/ from the University of Toronto.
Misarticulations of the /s/ are, in fact, common and often result from non-alveolar tongue position even in first-language speakers of English (McAuliffe & Cornwell, 2008; Mowrer & Sundstrom, 1988; Stephens & Daniloff, 1974). Indeed, while the /s/ is the fourth most commonly used consonant in English, it ranks as one of the most frequently misarticulated (Mowrer & Sundstrom, 1988). In pediatric populations of English first-language speakers, misarticulation of the /s/ often defies correction efforts (Daniloff, 1980). Causes of misarticulation have been explored, including delayed closure of the hard palate (Soderpalm & Lohmander-Agerskov, 1992), palate shape (Brunner, 2009), unusual pressure on the tongue when swallowing (McGlone & Proffit, 1973) and hearing loss (Ryalls, 1991). Misarticulation can produce a whistle or a blunted sound (Stephens & Daniloff, 1974) as well as lateral or interdental lisps (McAuliffe & Cornwell, 2008).

Research has suggested that this problem may result from difficulty in learning how to form a sound produced by forcing turbulent air through a precisely restricted aperture (Daniloff, 1980; Li et al., 2009). A study of consonant acquisition in English and Japanese children found that while children master stop consonants and glides early, the liquids and fricatives, including the /s/, tend to be acquired much later. By age three, only 62 percent of English-speaking children and 25 percent of Japanese-speaking children master the /s/ sound in their respective languages (Li et al., 2009).

For EFL and ESL instructors, the inability of many first-language speakers of Japanese to distinguish between the /s/ the /ʃ/ sounds ranks as one of the most troublesome consonant distinctions, just after the difficulty of perceiving and producing the difference between the English /l/ and /r/ (Nogita, 2010). In texts used to teach English to first-language speakers of Japanese, the /s/ sound is introduced as a voiceless
alveolar fricative. Indeed, books on Japanese phonology characterize the Japanese /s/ the same way, as alveolar (Vance, 1987). If first-language speakers of Japanese appear to be struggling with the distinction between the /s/ and the /ʃ/ sounds, English pronunciation-remediation guides emphasize the importance of mastering the alveolar tongue placement (Nogita, 2010). Nonetheless, difficulties persist (Lambacher, Martens, Nelson & Berman, 2010; Nakayama & Yamaguchi, 2003; Nogita, 2010).

Anne Crescini (personal communication, February 29, 2012), an EFL instructor in Japan, noted problems she encountered teaching this distinction to first-language speakers of Japanese the /ʃ/ of she when the students could clearly pronounce the /ʃ/ of the Japanese word, sushi, which she regards as the same sound. She also found it perplexing that her students would sometimes pronounce the word sea as she, clearly illustrating their ability to pronounce both sounds.

Nogita (2010) calls into question the common assumption that the confusion between the /s/ and the /ʃ/ result from the fact that the sounds are non-contrastive in Japanese. To the contrary, /s/ and /ʃ/ are contrastive before /a/, /u/ and /o/ (Helpern, 2012; Nogita, 2011). Examples cited by Nogita (2010) include [kasa], ‘umbrella’ and [kaʃa] ‘freight train’; [kasu], ‘sediment’ and [kaʃu], ‘singer’; [kaso], ‘depopulation’ and [kaʃo], ‘spot.’ While the /ʃ/ is often a non-contrastive allophone before the high-front vowel /i/, and no minimal pairs exist in the “central areas of the lexicon,” the /si/ and the /fi/ do appear contrastive in “peripheral lexical items, such as some cultural dialects and names” (Nogita, 2010, p. 15).

Nogita (2010) tested the ability of Japanese speakers to perceive the difference between the /si/ and the /fi/ by having a Japanese speaker trained as an announcer read
two utterances in Japanese, “A singer called Yosshî” and “A singer called Yossî,” to a group of 205 Japanese speakers, 93 of them monolingual speakers of Japanese; all of the participants could identify the distinction (Nogita 2010). Nogita (2010) had the same participants read a script of nonsense terms written in Katakana, the Japanese syllabary, that included the /si/ and /ʃi/. The scripts included the English tongue twister “She sells seashells by the seashore,” containing the /si/ and the /ʃi/ but replacing other troublesome consonant distinctions, the /ð/ and /z/, and /ɾ/ and /l/. The participants were told the utterance was Japanese, and none appeared to be familiar with the English tongue twister. The participants read from the script. One American first-language speaker of English and three first-language speakers of Japanese judges listened to the recordings and determined that all 205 speakers, regardless of gender or age, produced a clear distinction between the /si/ and the /ʃi/ (Nogita 2010). By contrast, studies have found that 80 percent of first-language speakers of Japanese reading this tongue twister in English orthography were unable to make this distinction consistently (Nogita, 2010). “Japanese speakers can distinguish between those syllables in Japanese, but when it comes to the English contexts, they somehow get confused,” he wrote (Nogita, 2010, p. 44).

Nogita (2010) hypothesized that the struggle of first-language speakers of Japanese have distinguishing the /s/ from /ʃ/ is actually induced by the characterization of the /s/ as an alveolar consonant. Indeed, Nogita (2010) produces data to support the hypothesis that avoiding the issue of alveolar placement altogether can produce better results in teaching the /s/ /ʃ/ distinction in English than teaching it as an alveolar consonant.
Significantly, several phonologists have suggested that an alternative non-alveolar articulation is a perfectly acceptable way to pronounce the English /s/. Calvert’s *Descriptive Phonetics* (Calvert, 1980) states that a minority of Americans produce an acceptable /s/ with the tip of the tongue approximating the lower incisors at the gum line, creating an /s/ sound that is perceived as both correct and native, a position he described as lingua-dental. In this position, the tip of the tongue is near the gum line of the lower incisors and the slightly grooved front of the tongue is raised toward the alveolar ridge, forming a narrow groove through which air is directed (Calvert, 1980). Raver-Lampman & Dossou (2011) investigated the prevalence of the acceptable alternative articulation of the /s/ sound in word-final position in English-speaking university students and found a majority of their sample of 50 speakers pronounced the /s/ with their tongue in contact with the lower incisors.

An EFL instructor in Japan maintained that Japanese speakers pronounce the Japanese /s/ with this placement routinely (Cairns, 1988). He asserted that when Japanese speakers pronounced the /s/ in Japanese the tongue is “almost always in contact with the teeth, and the lower teeth, if my observations are sound” (Cairns, 1988, p. 50).

While Japanese phonology texts describe the Japanese /s/ as an alveolar fricative, no research found in a review of literature has examined the articulation of the Japanese /s/ with technology that can show how the tongue moves in the oral cavity when Japanese speakers pronounce the /s/ in their own language. If the Japanese pronounce the /s/ in their own language in the manner Cairns (1988) described, a position which is also an acceptable alternative pronunciation of the English /s/ (Calvert, 1988; Raver-Lampman &
Dossou, 2011), problems resulting from the emphasis to first-language speakers of Japanese on the alveolar placement of the English /s/ (Nogita, 2010) could be explained.

This study tested the hypothesis that Japanese routinely pronounce the /s/ in their own language with the apex of the tongue projected downward, the same placement posited by Calvert (1980) of the English /s/ and observed by Cairns (1988) for the Japanese /s/.

The tongue position used in articulating the English alveolar /s/ can vary, with the tongue position ranging from the top of the upper incisors to just behind the alveolar ridge; the tongue position of the mandibular /s/ is likely to show such variance as well. Nothing in the literature suggests an alternative articulation of the English or Japanese /s/ in the interdental position. Thus, for purposes of this study, articulation of the /s/ with the tongue’s apex projecting upward toward the roof of the oral cavity will henceforth be called alveolar; the position with the tongue’s apex projecting downward toward the mandible will henceforth be called mandibular.

To test this hypothesis, it was necessary to find a means to capture the placement of the tongue as first-language speakers of Japanese pronounce utterances that contain the sound /s/ in their own language.

1.2. REVIEW OF MODALITIES TO CAPTURE DYNAMIC ARTICULATION

Techniques that have been used to determine tongue positions during speech include static palatography (Ladefoged, 2003), electropalatography (Cheng et al., 2007; McAuliffe & Cornwell, 2008), magnetic resonance imaging (MRI) (Takano & Honda, 2007), X-ray (Ladefoged, 2012; Mielke, Baker, Archangeli, & Racy, 2005) and
ultrasound (Davidson, 2006; Gick, 2008; Mielke et al., 2005; Parthasarathy, Stone, & Prince, 2005).

Each has its limitations. Static palatography initially employed chocolate powder dusted on the palate that allowed linguists to determine tongue position by observing where the powder was disturbed during articulation (Ladefoged, 2003). Static palatography has been modified, but still requires the use of a palate-dusting medium (Ladefoged, 2003). The advantage of palatography is that it is inexpensive and requires no specialized equipment. Its successor, eletropalatography, uses an electronics-studded palate liner to record the tongue position electronically (Cheng et al., 2007). In both cases, however, these techniques record where the tongue comes into contact with the palate; both are silent on articulations in which the tongue touches the lower teeth.

Magnetic resonance imaging (MRI) can produce accurate images of the tongue (Takano & Honda, 2007), but MRI devices are expensive, and they require the participant to be examined lying down. In addition, these devices/instruments suffer from poor temporal resolution, making it impractical for recording movement of the tongue (Mielke et al., 2005).

X-rays can record tongue movement in real time with great clarity, including images of the teeth and other surrounding structures (see Figure 2), but the modality is immobile and the radiation required to create the image is toxic to participants (Stone, 2005).
Diagnostic ultrasound, which relies on sound waves to create images, has been used in the medical field since the late 1950s. Ultrasound captures images by generating ultra-high frequency sound waves projected from a transmitter-receiver, or transducer. Those waves are reflected back to the transducer by various physical structures and a computer produces an image by calculating the time it takes to reflect from a surface at the speed of sound (Stone, 2005). Ultrasound produces no toxic radiation and can capture tongue movement well. Ultrasound is starting to be utilized as a linguistic research tool as it grows more powerful and portable (Gick, 2002).

While ultrasound does have limitations outlined below, it seemed the best modality available to capture the tongue movement for the present study.
CHAPTER II

METHOD AND PROCEDURE

2.1. USE OF ULTRASOUND

Ultrasound devices render images from transducers designed to project an image of a specific surface at a specific depth (Mielke et al., 2005). Ultrasound devices can output videos at 29.97 frames per second. While the technology has been used in linguistics research, until recently it, too, was static and often required the use of a special room and an immobile device (Gick, 2002). Recent developments have produced ultrasound machines that are portable and more appropriate for linguistic fieldwork (Gick, 2008; Mielke et al., 2005; Stone, 2005).

To image the movement of the tongue, an ultrasound device uses a palm-sized transducer (see Figure 3) in the mandibular window, under the jaw and running between neck and the chin, bounded by the mandible.

![Figure 3. Mindray C5 2S Transducer.](image)

Ultra-high frequency sound waves generated by the transducer project upward and reflect
back from where the tongue comes in contact with intraoral air, creating an image from
the back of the oral cavity to the chin (Stone, 2005; Gick, 2002).

2.2. LIMITATIONS OF ULTRASOUND

Ultrasound does have limitations for linguistics research, however. The image is
relatively grainy and, unlike X-ray, an ultrasound image of the tongue does produce
images of the surrounding structures such as the teeth and the palate. Moreover, because
ultrasonic waves cannot penetrate bone, an acoustic shadow produced by the mandible at
the chin partially obscures the tongue’s apex and a shadow produced by the hyoid bone
partially obscures the tongue root (see Figure 4).

2.3. CREATION OF A POINT OF REFERENCE

An oval border along the surface of the tongue in the neutral positions was created
using an image-processing program called ImageJ. This program separated the video into
stacks of tagged image file format or tiff images, each representing a single frame of the
video. To create a reference point, a frame was extracted showing a participant’s tongue
in a neutral position, neither touching the roof of the mouth nor projecting to the floor of
the oral cavity; in this position, the apex of the tongue is generally pointing to where the
incisors meet (Hall, 2012). ImageJ includes a polygon function that allows one to drop
points along the surface of an image, and then finish the shape by placing the end point
on top of the beginning point, creating a polygon. ImageJ’s spline-creating function then
transforms this polygon onto a smooth oval surrounding the image. One can save the
image from the single image to all the frames in the video, providing a consistent
reference point for analysis.

**Figure 4.** Annotated Illustration of an Ultrasound Image of the Oral Cavity Used in This Study with the Tongue in the Neutral Position, Neither Touching the Roof of the Mouth nor Depressed to the Floor of the Oral Cavity.

2.4. ACOUSTIC SHADOW AND ATTENUATED REFLECTION

On ultrasonic images, the acoustic shadow produced by the chin grows wider toward the roof of the oral cavity, obscuring more of the tongue’s apex. An illustration of this can be seen in Figure 5 below, which displays an alveolar /s/ producer from the study sample pronouncing the /s/ of *kudasai* (see Figure 5).
The amount of the tongue’s apex obscured is a product of a participant’s anatomy and cannot be corrected through adjustments of the ultrasound device or transducer (Stone, 2005). If the tongue’s apex is too obscured, it can render some data unusable. Data were discarded if the acoustic shadow obscured the apex of the tongue to the point that upward or downward projection could not be determined.

Another limitation is that as the tongue’s surface approaches the roof of the oral cavity, the pocket of intraoral air along the top of the tongue that creates the ultrasonic reflection grows smaller or vanishes, attenuating or eliminating the tongue’s ultrasonic image (Stone, 2005). This is illustrated with the articulation of the velar plosive /k/, in which the tongue presses against the soft palate (see Figure 6).
In Figure 6, the dorsum of the tongue vanishes as it comes into contact with the soft palate to form the [k] of *kudasai*.

Awareness of these ultrasonic limitations as well as the articulatory characteristics of phonemes can allow one to draw conclusions about the articulatory placement of the tongue (Stone, 2005). For the purposes of this study, the upward thrust of the tongue is consistent with the alveolar placement of the /s/ sound and thus deemed to be alveolar; the downward projection of the tongue is consistent with the articulation at the back of the lower incisors, the articulation described by both Calvert (1980) and Cairns (1988), and is deemed mandibular.
2.5. SYNCHRONIZATION OF VIDEO TO AUDIO TRACK

Ultrasound devices record silent video images that have to be synchronized to an audio track recording the sounds being uttered. Synchronization of the silent ultrasound image to a separately recorded audio track presents difficulties (Stone, 2005). In research requiring precise data on distance or contact points between the oral structures and the tongue, researchers must obtain an ultrasonic image of the oral cavity and add a third track, a video showing the participant during the ultrasound session with a fixed device that allowed research to mathematically factor in head movement. Such techniques require head stabilizers, such as dental chairs, or other equipment to minimize or measure any head movement. Adding a third data video track increases the complexity of synchronization. Magnuson (2011) proposed testing synchronization of all three tracks by placing a water balloon on the ultrasound transducer, pulling on the latex ring and snapping it. The sloshing of the water in the balloon is visible on the ultrasound image; the snapping sound is recorded on the audio track and the visual image of the latex ring being snapped is captured on the video recording of the participant’s session. Hueber, Cholet, Denby & Stone (2008) checked alignment by tapping a hammer against a bottle of ultrasound gel so that a droplet appears immediately on the ultrasound image, the sound of the hammer is recorded on the audio track and the hammer is seen striking the bottle on the video recording. Miller & Finch (2011) used the ultrasound machine’s internal Digital Imaging and Communications in Medicine (DICOM) format, which produces files that measure many gigabytes, to generate a 124-frame-per-second ultrasound image to study the IsiXhosa alveolar click, a phoneme far too rapid to be captured on the standard 29.97-frame-per-second rate. Synchronization software was
checked with a device that produced a 3-millisecond ultrasonic disturbance, a simultaneous buzzer recorded by the audio track and a light emission captured by the video recording the session. While such precision may be required for the visualization of extremely rapid consonant sounds such as the IsiXhosa alveolar click, Miller & Finch (2011) acknowledge that the standard 29.97 frames per second is adequate to produce usable images of most vowels and fricatives.

For this study, the alveolar and mandibular designations were determined by the direction of the tongue's apex, which could be captured without elaborate stabilizers or a third video track. That left just two tracks, the silent ultrasound video and an audio track. A novel method of synchronization was employed, a distinct and rapid feature included in all utterances, the release of the velar plosive /k/. On the ultrasound videos, the /k/ sound produces visual image that is different from that produced by the other sounds in the utterances. On the ultrasound image, the formation of the /k/ is preceded by the rise from the neutral position (see Figure 7) of the tongue’s dorsum to the soft palate (see Figure 8). The sound of the /k/ occurs as the tongue pulls away from the palate rapidly, a movement captured in a single frame of the video, which records at 29.97 frames per second (see Figure 9). This frame is aligned with the sound of the release of the /k/ recorded on the audio track. This provides synchronization down to a single video frame.

2.6. ULTRASOUND DEVICE EMPLOYED

This research was conducted with a portable device manufactured by Mindray, a manufacturer of ultrasound devices frequently used in medical diagnostics. The Mindray M-5 has a laptop-sized imaging system with a variety of transducers used for various
Figure 7. Participant 3, Frame 161. Tongue in the Neutral Position Before Uttering the Japanese Word *Kasa*.

Figure 8. Participant 3, Frame 174. The Dorsum of the Tongue Rises Toward the Soft Palate to Form the [k] of *Kasa*.
Figure 9. Participant 3, Frame 176. The Tongue Descends Rapidly from the Soft Palate at the release of the [k] in Kasa, Creating the /k/ Sound.

diagnostic applications. The transducer best for producing an image of the tongue was the convex C5-2s transducer shown in Figure 2. The transducer both generates and receives ultra-high frequency vibrations, in this case between 2.5 and 5 million oscillations per second, the appropriate range for creating an image of the tongue’s surface. The convex shape allows the transducer to be nestled under the jaw, running from the chin to the neck, bounded by the mandible.

2.7. SONOGRAPHER

Dr. Felicia M. Toreno, Ph.D., RDMS, RDCS, RVT, Director of Diagnostic Medical Sonography at Tidewater Community College, provided, operated and calibrated the device and helped provide background necessary to read the images. The audio was recorded with a MicroTrack II a two-channel, 24-bit, 96-kilohertz digital recorder that
produces files of sufficient quality to be edited by PRAAT acoustic analysis software (Boresma & Weenik, 2010).

2.8. STUDY MATERIALS

This study required the creation of documents in Japanese that could be used to test the pronunciation of the Japanese /s/ in word-initial and word-medial positions. Words and phrases were chosen from Japanese instruction texts for English speakers (Jorden, 1963; Pegasus Language Services, 1987), opting for extremely simple words that Japanese speakers could pronounce easily. A professional Japanese translator was hired to consult on word and sentence choices and also to translate the selected words and phrase into the two common Japanese writing systems, Hiragana and Kanji (see Appendix 1). Both scripts were necessary because Japanese has multiple writing systems and some Japanese speakers can read in one but not the other. The word list also included instructions in Japanese.

Because ultrasound devices can record only short video segments, each participant had to repeat multiple utterances to capture images of the /s/ sound in various phonemic environments and in both word-initial and word medial position. Three sets of utterances were developed, the first containing three words, the second containing four words and the third consisting of one short sentence. Each utterance contained the /k/, which was used to synchronize the ultrasound video with the audio track (see Table 1).

Japanese is a mora-timed language and two vowels in a word such as kiite each bear the same time, creating a “long” vowel (Vance, 1987). Although the Japanese
language generally does not contain diphthongs, Vance (1987) posits that a word that ends with an [a] followed by a high-front vowel [i], such as the [ai] of kudasai, is pronounced as a diphthong. The audio data gathered from the Japanese-speaking participants in this study confirmed that the final [ai] of kudasai is, in fact, pronounced as a diphthong. An /a/ before the /n/ in words such as sansuu is nasalized (Vance, 1987), producing the sound [ã], a phenomenon recorded in the data.

The two lists of words and the short sentence permitted the examination of the /s/ sound in the following 11 environments (see Table 2).

Table 2
Phonemic Environments of the /s/ in Utterances Used in the Study.
2.9. SETTING

The study took place at the Newport News Japanese School, a Saturday school for children ranging from kindergarteners to 9th graders who study Japanese to supplement their education in American schools. The Newport News Japanese school is attended by approximately 70 students and employs 15 teachers whose first language is Japanese. The study was conducted over three weekends. The research team consisted of the principal investigator, the sonographer and the translator. Participants were volunteers who were teachers or parents.

Research took place in a room set aside for the team. Participants entered the room, often with their children. As the study focused on the visual image of the video produced by the ultrasound device, no effort was made to sound-shield the room in the manner required for acoustic research. Participants completed an informed consent form translated into Japanese. Study participants then completed a demographic survey translated into Japanese (see Appendix 2). All documents were numbered sequentially, from 1 to 47, the total number of participants who completed the study.

2.1.0. DATA SELECTION CRITERIA AND DEMOGRAPHIC CHARACTERISTICS OF PARTICIPANTS

The principal of the Newport News Japanese school assisted in recruiting both teachers and parents for the study using the selection criteria which had been provided: Participants had to be raised in Japan and had to speak Japanese as their primary
language until age 10. Because this was a convenience sample, it was impossible to balance for age, gender or other demographic characteristics.

The 20 participants who produced video and audio data useable for the study, however, came from throughout Japan, from the far southern islands to the northernmost prefecture. The sample included a total of 17 females and 3 males, 10 in their 30s, 8 in their 40s and 2 in their 50s. All self-identified has having spoken primarily Japanese until they were 10 years old (see Table 3). Most had very limited English.

Table 3
Demographic Survey.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Gender</th>
<th>City of Birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40s</td>
<td>F</td>
<td>Nagano</td>
</tr>
<tr>
<td>2</td>
<td>50s</td>
<td>F</td>
<td>Tokyo</td>
</tr>
<tr>
<td>3</td>
<td>30s</td>
<td>F</td>
<td>Okinawa</td>
</tr>
<tr>
<td>4</td>
<td>30s</td>
<td>F</td>
<td>Nara</td>
</tr>
<tr>
<td>5</td>
<td>50s</td>
<td>F</td>
<td>Akita</td>
</tr>
<tr>
<td>6</td>
<td>40s</td>
<td>M</td>
<td>Kanagawa</td>
</tr>
<tr>
<td>7</td>
<td>30s</td>
<td>M</td>
<td>Hakkaido</td>
</tr>
<tr>
<td>8</td>
<td>30s</td>
<td>F</td>
<td>Miyazaki</td>
</tr>
<tr>
<td>9</td>
<td>40s</td>
<td>F</td>
<td>Fukuoka</td>
</tr>
<tr>
<td>10</td>
<td>30s</td>
<td>F</td>
<td>Hyogo</td>
</tr>
<tr>
<td>11</td>
<td>30s</td>
<td>F</td>
<td>Fukuoka</td>
</tr>
<tr>
<td>12</td>
<td>30s</td>
<td>F</td>
<td>Saitama</td>
</tr>
<tr>
<td>13</td>
<td>40s</td>
<td>F</td>
<td>Okinawa</td>
</tr>
<tr>
<td>14</td>
<td>40s</td>
<td>F</td>
<td>Saitama</td>
</tr>
<tr>
<td>15</td>
<td>40s</td>
<td>F</td>
<td>Tokyo</td>
</tr>
<tr>
<td>16</td>
<td>40s</td>
<td>F</td>
<td>Kanagawa</td>
</tr>
<tr>
<td>17</td>
<td>40s</td>
<td>F</td>
<td>Kanagawa</td>
</tr>
<tr>
<td>18</td>
<td>30s</td>
<td>M</td>
<td>Shimane</td>
</tr>
<tr>
<td>19</td>
<td>30s</td>
<td>F</td>
<td>Chiba</td>
</tr>
<tr>
<td>20</td>
<td>30s</td>
<td>F</td>
<td>Nagasaki</td>
</tr>
</tbody>
</table>
While ultrasounds were completed on 47 participants, not all the data were complete or suitable for analysis. Due to an unanticipated video preset on the ultrasound device, during the first session, only three seconds of each ultrasound video were recorded, too short to fully capture any of the utterances, rendering those 15 ultrasound videos unsuitable for analysis. Nonetheless, those videos served as a confirmation that the ultrasound adequately captured tongue movement without employing a stabilizing device and that video and audio of each utterance could be successfully aligned using the release of the velar plosive /k/. In the two subsequent data-collection sessions, the ultrasound device was manually configured to record videos lasting eight seconds. Nonetheless, the eight-second video segments failed to capture the complete utterances of 6 participants. With an additional 4 participants, the video contrast was too low to provide a reliable reading of the tongue position. In the case of 2 participants, the tongue’s apex was too obscured by the acoustic shadow to determine its direction.

2.1.1. STUDY PROCEDURE

The laptop-sized ultrasound device and the digital voice recorder were placed on a table. When a participant was seated, the principal investigator turned on the digital voice recorder and stated the number of the participant. Because of the limited English-speaking capabilities of most participants, the translator demonstrated how to hold the transducer and explained that participants would read aloud from three lists of Japanese words.

The principal investigator started the Microtreck II digital voice recorder at the beginning of the session and called out the number of the participant. The sonographer
instructed each participant to hold the transducer in place and move the tongue up and
down, while adjusting the position of the transducer to increase the resolution of the
tongue’s surface. When the sonographer was satisfied with the resolution, the participants
were instructed to read two lists of words and one Japanese sentence. Again, because of
an eight-second limit on the length of the ultrasound videos, both word lists and the short
sentence had to be recorded separately.

To avoid head movement during the reading, each list of words was printed in 48-
point font on 8-by-11 sheet of paper placed in landscape orientation on a desktop paper
holder in front of the participant. After the first list of three words was read, the first list
was removed to reveal the second word list; the second list was removed to reveal the
short sentence. If head or transducer movement caused the image to degrade, the
sonographer asked the participant to repeat a set of words or the sentence. Each
participant generated a minimum of four ultrasound files, including a video showing
tongue movement, the two sets words and the sentence.

The principal investigator ended the audio recording after the entire sequence was
completed, creating one long audio file to be segmented and aligned with the multiple
ultrasound video files. Each video file contained an exemplar of the velar plosive /k/.
After each participant’s session, the Japanese translator translated and numbered the
demographic survey to correspond with the number of the participant.

2.1.2. PRESERVATION OF THE DATA

All ultrasonic video files were exported in audio video interleave (AVI) format
and saved on a LaCie 9230 one-terabyte external hard drive configured to accept files
from Windows, the operating system of the Mindray ultrasound device, and Macintosh, the operating system of the computer used to analyze the video files. The sound files produced by the Microteck II were saved in audio interchange file format (AIFF) in the same folder. In total, the ultrasound video and audio files were saved in 47 folders, corresponding to the number of a participant.
CHAPTER III
ANALYSIS OF THE DATA

3.1. ANALYSIS OF THE ULTRASONIC IMAGES AND SOUND FILES

Apple's iMovie 11, operating on a MacBook, was used to combine the video and
the audio of each utterance aligning the sound files to the video files at the velar
consonant /k/, which appeared in each list of words and also appeared three times in the
short sentence. As had been described, the ultrasound images were grainy (Gick, 2002;
Mielke et al., 2005; Stone, 2005), with the tongue's surface the most salient feature. Also,
as predicted (Gick, 2002; Mielke et al., 2005), the tongue's apex was obscured by the
acoustic shadow created by the mandible at the chin.

The folder for each participant contained three ultrasound videos numbered 1, 2
and 3, and one long audio file. Using PRAAT (Boresma & Weenik, 2010), the single
audio file was segmented to correspond to the ultrasound videos of the participants
speaking the three utterances: the first audio file segment consisted of the participant
speaking the words sansuu, kasa and sensei; the second segment consisted of the
participant speaking the words sumoo, miso and soko desu; and the third segment
consisted of the participant speaking the short sentence, Sasaki-san, soko ni suwatte
kudasai.

Using iMovie, the researcher imported the ultrasound video of a participant
speaking the first utterance as well as the respective audio file. The sound of the /k/ on
the audio file was aligned with the image of its release. In cases where the tongue was
quiescent before and after speaking, alignment was confirmed by pauses at the beginning
and end of the utterances. In the short sentence, *Sasaki-san, soko ni suwatte kudasai*, the three /k/ sounds lined up, confirming synchrony.

Reading the ultrasound images produced in the study required a knowledge of the articulation of all the phonemes in the utterances (Stone, 2005). The analysis also had to take into consideration several allophonic articulations, including the following:

- Participants in the study produced the [o] by dropping the apex of their tongues to the floor of the oral cavity in the word-final position (see Figure 10).

![Figure 10. Participant 1, Frame 155. Pronouncing the Word-Final [o] with the Tongue on the Floor of the Oral Cavity.](image)

- Participants also placed the apex of the tongue against the floor of the oral cavity when participants pronounced the [α] preceded by the labio-velar central approximant [w].
• Participants pronounced the [t] and the [d] with the apex of the tongue against the top teeth, an articulation described by Vance (1987).

• Although all Japanese syllables must end in a vowel, a final /w/ is voiceless (Vance, 1987). In the data, participants did not voice the final /w/ in the word desu, but they did move their tongue to the position required to pronounce it.

ImageJ image-processing software allowed frame-by-frame contrast adjustment. ImageJ also allowed the researcher to create the oval border tracing participant’s tongue in a neutral position, a common reference point that increases the visual salience of the tongue’s movement upward or downward.

Below is a frame-by-frame analysis of participant 9, a mandibular producer, uttering the short sentence Sasaki-san, soko ni suwatte kudasai, 'Mr. Sasaki, please sit down here' (see Figures 11-28).

Figure 11: Participant 9 (Mandibular Producer), Frame 1. Tongue in the Neutral Position.
Figure 12. Participant 9, Frame 71. Tip of Tongue Descends to Form First [s] of Sasaki.

Figure 13. Participant 9, Frame 79. Tongue Drops Further to Form the Second [s] of Sasaki, Formed after the Vowel [a]
Figure 14. Participant 9, Frame 90. Dorsum of the Tongue Vanishes as it Makes Contact with the Soft Palate to Articulate the [k] of Sasaki.

Figure 15. Participant 9, Frame 98. Back of Tongue Drops, Apex Rises to Form the High Front Vowel [i] of Sasaki.
Figure 16. Participant 9, Frame 112. Tongue Tip Descends from the [i] to Form the [s] of San.

Figure 17. Participant 9, Frame 128. Tongue Tip Moves Upward Through the Nasalized [ä] and Pushes Forward and Upward to Form the [n] of San.
Figure 18. Participant 9, Frame 150. Tongue Tip Descends from the [n] to Form the [s] and Coarticulated [o] of Soko.

Figure 19. Participant 9, Frame 161. The Dorsum of the Tongue Rises and Vanishes as it Makes Contact with the Soft Palate to Form the [k].
**Figure 20.** Participant 9, Frame 166. The Tongue's Apex Pushes Forward and Upward to Form the \([n]\) of \(Ni\).

**Figure 21.** Participant 9, Frame 172. The Tongue Rises to Form the High Front Vowel \([i]\) of \(Ni\).
Figure 22. Participant 9, Frame 185. The Tongue Tip Descends to Form the [s] of Suwatte.

Figure 23. Participant 9, Frame 187. The Tongue Descends Further to Form the [wa] of Suwatte Toward the Floor of the Oral Cavity.
Figure 24. Participant 9, Frame 201. The Tongue Tip Rises to the Back of the Front Teeth to Form the Dentalized Articulation of the [t] of Suwatte.

Figure 25. Participant 9, Frame 204. The Dorsum of the Tongue Vanishes as it Comes into Contact with the Soft Palate to Form the [k] of Kudasai.
Figure 26. Participant 9, Frame 212. The Tongue’s Apex Rises Toward the Top Teeth Forming the Lamino-Alveolar [d] of Kudasai.

Figure 27. Participant 9, Frame 220. The Tongue Descends to Form the [s] of Kudasai.
3.2. INSUFFICIENT DATA ON /S/ BETWEEN ALVEOLAR AND MANDIBULAR VOWEL

While the literature does not describe it, the participants in this study pronounced the word-final /o/ in a mandibular position (see Figure 29), with the tongue against the floor of the oral cavity.

In the case of the word miso, mandibular /o/ occurs immediately after the high front vowel /i/. The /s/ falls between these vowels. As Hume (1994) observed, the placement of the high front /i/ is often indistinguishable from the placement of the alveolar /s/. In this case, /o/ pronounced with a mandibular tongue placement is very similar to the placement of the mandibular /s/. As measured on PRAAT (Boresma & Weenik, 2010), the elapsed time between the pronunciation of the /i/ and the /o/ averaged .16 seconds. With the video recording data at a rate of 29.97 frames per second, data
recorded in those .16 seconds is limited to a few frames. While frames do contain images that could be interpreted as supporting the hypothesis that Japanese speakers commonly use a mandibular articulation for the /s/, the data produced in the interval is extremely limited. To be conservative, data from the word *miso* were eliminated from the study.

*Figure 29. Participant 1, Frame 155. Mandibular Producer Pronouncing the [o] of *Miso* with the Placement Used by Participants When the Vowel Appeared in Word-Final Position.*

That left a total of 20 participants and 14 tokens per participant. Those 14 tokens included 8 word-initial, 6 word-medial for a total of 280 tokens of the /s/.
CHAPTER IV

RESULTS

The study found that in 17 of 20 participants (85 percent) the apex of the tongue projected downward consistent with the mandibular position posited by Calvert (1980) and Cairns (1988); in 3 of 20 participants (15 percent) the apex of the tongue rose toward the roof of the oral cavity when producing the /s/ sound, consistent with the alveolar placement of the /s/ listed on IPA charts and promulgated in pronunciation guides (see Graph 1).

![Graph 1](image-url)

*Graph 1. Results of Study by Participants Using IPA-Described Alveolar Articulation or the Acceptable Alternative Mandibular Articulation.*

The 17 participants who pronounced the /s/ in mandibular position did so in both word-initial and word-medial position; the alveolar producers showed the same consistency, with the alveolar articulation of the /s/ remaining alveolar word initially and word medially (see Table 4).
Participants who pronounced the /s/ in the alveolar position did show vowel effects, an altered articulation of the /s/ in specific phonemic environments. With the standard articulation by alveolar producers, the /s/ is pronounced with the body of the tongue level, close to a neutral position, while the apex projects upward. This is illustrated in Figure 30, the /s/ articulated in the word *kasa*.

Table 4  
Breakdown of /s/ Sounds in Word-initial and Word-Medial Positions by Participant.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Mandibular - initial</th>
<th>Mandibular - medial</th>
<th>Alveolar - initial</th>
<th>Alveolar - medial</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>20</td>
<td>8</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>136</td>
<td>102</td>
<td>24</td>
<td>18</td>
</tr>
</tbody>
</table>
In Figure 31, in the /s/ pronounced by an alveolar producer, the body of the tongue remains level, within the oval border surrounding the tongue in the neutral position, while the apex rises toward the alveolar ridge to form the /s/ of *kasa*.

![Image](image.png)

*Figure 30. Participant 8 (Alveolar Producer), Frame 182.*

This contrasts with the tongue position of both the vowel and the consonant in alveolar /s/ producers pronouncing the /s/ before the low front /a/ followed by the high front /i/ in the diphthong /ai/ of *kudasai*. In this case, body and dorsum of the tongue rise toward the soft palate to form the /s/ and the coarticulated /a/ and lower slightly to eliminate the turbulent air that creates the fricative sound (see Figures 31 and 32).

In Figure 32, both the body and apex of the tongue rise above the neutral position toward the palate to produce the /s/ in of *kudasai*. A portion of the apex is obscured by the acoustic shadow.
In Figure 32, the body and the apex of the tongue lower slightly to eliminate the turbulence required to generate the fricative sound of the /s/ and to produce the final high-front vowel /i/ of kudasai. This change comports with the finding of Lee (1997) that a vowel can have an effect on the consonant preceding it. Vowel effects could also be
observed with mandibular producers in producing the same diphthong, the final /ai/ of *kudasai*.

In mandibular producers, the tongue’s apex descends lower into the oral cavity when pronouncing the /s/ than it does in other before other vowels, creating an extremely salient visual difference between the alveolar /s/ and the mandibular /s/, although this did not create any discernable auditory change (see Figures 33 and 34).

![Alveolar producer](image1.jpg) ![Mandibular producer](image2.jpg)

*Figure 33. Participant 3 (Alveolar Producer), Frame 268. The Body and Apex of the Tongue Rise Toward the Palate to Produce the [s] of *Kudasai*.  
Figure 34. Participant 9 (Mandibular Producer), Frame 220. Tongue Descends Toward the Floor of the Oral Cavity to Form the [s] of *Kudasai*.  

An extremely salient articulatory contrast between the alveolar producers and the mandibular producers also occurred in the production of the word-final vowel in word *kudasai* (see Figures 35 and 36).

This contrast in both the /s/ and the word-final vowel also occurs when the /s/ is followed by long close mid-front word-final /e:/ of *sensei*. 
In summary, the findings confirm that the hypothesis that most first-language speakers of Japanese in this study pronounced the /s/ in the mandibular position, an acceptable alternative position employed by many first-language speakers of English.

*Figure 35.* Participant 3 (Alveolar Producer), Frame 283. The Tongue Lowers from the Alveolar Ridge to Eliminate the Turbulent Sound of the Fricative to Pronounce the High Front Final Diphthong of *Kudasai.*

*Figure 36.* Participant 9 (Mandibular Producer), Frame 235. Tongue Ascends to Form the High Front Final Vowel of the Diphthong [ai] of *Kudasai.* The Place of Production of the Vocalic [ai] in the Mandibular Producer Differs from the Placement by the Alveolar Producer.
CHAPTER V
DISCUSSION AND CONCLUSION

5.1. DISCUSSION AND IMPLICATIONS FOR INSTRUCTION

This study partially validates the observation by Cairns (1988) that Japanese speakers often produce the /s/ with their tongues touching their lower incisors. Cairns (1988), however, showed no awareness of the observation by Calvert (1980) that the first-language speakers of English produce the /s/ using the mandibular articulation he observed in his Japanese students. Instead, Cairns (1988) concluded from his observation that the Japanese /s/ he observed was incorrect and produced a defective English /s/.

Cairns urged ESL and EFL teachers to focus on teaching the English /s/ in the alveolar position, thus "reducing dependence on the mother tongue" (Cairns, 1988, p. 53). Although Cairns' (1988) was not widely cited or acknowledged, pronunciation guides on the Internet and sold in Japan routinely emphasize the importance of the alveolar articulation (Becker, 2010; Nogita, 2010).

It goes without saying that the purpose of ESL and EFL instruction is to convey to students the most efficient way to master not only the syntax and lexis of a language but the proper method to pronounce the sounds of the language, including what has been described as the voiceless alveolar fricative /s/. Continuing the practice of instructing Japanese ESL and EFL students to use the alveolar would be easy to justify if it were successful, but correcting the problems with the /s/ using such techniques has not shown itself to be effective.

The research of Nogita (2010) strongly suggests that teaching the alveolar position may actually increase errors. Nogita (2010) found that the most effective
approach to teaching first-language Japanese speakers to pronounce the difference between the /s/ and the /ʃ/ was to avoid teaching them the alveolar placement. Nogita (2010) demonstrated the ability to improve the ability to perceive and produce the difference in a 40-minute teaching session focusing on English phonics and orthographic representation of the /s/ and the /ʃ/, ignoring the articulation altogether. “Pronunciation textbooks should focus more on the basic phonological rules, or when to pronounce each phoneme in a word, not how to pronounce it,” he concluded (Nogita, 2010, p. 122).

The present study goes further in its finding that the articulation of the Japanese /s/ used by 85 percent of the Japanese participants in this study is similar or identical to the mandibular articulation described by Calvert (1980); this mandibular articulation was also found in a majority of participants in a study of the articulation of /s/ of first language speakers of English (Raver-Lampman & Dossou, 2011). The functional effect of Nogita’s (2010) avoidance of teaching the alveolar articulation is that it allowed Japanese speakers to continue to use the articulation they employed in their own language. In other words, the problem was not teaching articulation in general, but in teaching an unfamiliar articulation of two sounds, the /s/ and the /ʃ/, that first-language speakers of Japanese had already mastered.

It’s important to note that this study does not suggest that ESL and EFL teachers need to intervene to correct problems in first-language speakers of Japanese whose English is functional and comprehensible. In truth, the alveolar placement may be emphasized early in the study of English, as the sounds of English are being introduced. If anything, the pronunciation-remediation and accent-reduction guides asserting the importance of the alveolar placement tend to agree with Cairns that the non-alveolar
English /s/ is aesthetically deficient (Cairns, 1988; Li, Edwards & Beckman, 2009; Nogita, 2010). The results of this study could explain why introducing the English /s/ as an alveolar consonant to first-language speakers of Japanese can induce errors and complicate an already daunting task, learning a new language.

In summary, these findings not only add to the literature, they also offer concrete guidance that ESL and EFL instructors can use to improve the teaching of this troublesome distinction to first-language speakers of Japanese.

5.2. LIMITATIONS

This research has limitations. First, the tongue’s apex in ultrasounds is not entirely visible making it impossible to state conclusively the contact point on the incisors. Second, researchers conducting such research must develop a working knowledge of ultrasound to interpret the images. Despite these limitations, this study demonstrated the ability to use a portable ultrasound device to capture the basic articulatory position of both consonants and vowels in the sample without employing a stabilizing device and using a phonemic feature of synchronize the audio and video tracks. In years past, determining the actual tongue position of people speaking any language was extraordinarily cumbersome. This study has demonstrated a method to use ultrasound in the field to answer basic phonemic questions.

5.3. FUTURE RESEARCH AND CONCLUSION

Without empirical testing, assumptions about articulatory position remain just that, assumptions. This study has contributed to the literature by showing how ultrasound
can be used in the field to support the hypothesis that Japanese use a mandibular articulation of the Japanese /s/. As ultrasonic technology becomes portable enough to be deployed in the field, efforts to confirm articulatory placements and their alternatives becomes increasingly feasible. Such research could begin in English and map English sounds and allophones including acceptable alternative placements. Such a phonemic map in English could be compared to articulatory maps of other languages to find cross-linguistic commonalities in the productions of vowels and consonants, potentially reducing difficulties teaching challenging sounds.

Paul Passy (1907), the linguist widely credited with creating both the IPA and much of the phonetic alphabet in use today, would certainly not approve of continuing to teach the IPA descriptions prescriptively. In his volume on French phonology (Passy, 1907), Passy writes that the French /s/ is pronounced “with the tip of the tongue against the lower front teeth” (Passy, 1907, p. 82), a mandibular production. He states that the acoustic difference between mandibular /s/ and the alveolar /s/ is “practically negligible” (Passy, 1907, p. 82).

Today, the tools are available, to allow future ultrasonic studies by researchers working independently to create a comprehensive articulatory map of sounds in English and other languages. Such a resource would offer a valuable contribution to the field of linguistics and could also benefit those who work daily in ESL and EFL classrooms.
REFERENCES


APPENDIX A

STUDY WORDS FOR ENGLISH SPEAKERS

INSTRUCTIONS: Below are two sets of words and a sentence. When instructed, read the first group of words. Pause after the first group and then read the second group when instructed to begin again. Pause after you read the second group and read the sentence when instructed to do so. You may be asked to reread if there are technical problems.

Group A

1) さんすう (算数) -- Sansuu (math)
2) かさ (傘) -- kasa (umbrella)
3) せんせい (先生) -- Sensei (teacher)

Group B:

1) すもう (相撲) -- Sumoo (sumo wrestling)
2) みそ (味噌) -- miso (soybean paste),
3) そこです -- soko desu (… is there)

Sentence

ささきさん、そこにすわってください。 （佐々木さん、そこに座ってください。）

Sasaki-san, soko ni suwatte kudasai  --  Mr. Sasaki, please sit down there
APPENDIX B

STUDY WORDS IN JAPANESE

研究で使用する単語（ひらがな・漢字の原稿）

方法の説明:
2つのグループに分けた単語と1つの文章が下記に示してあります。指示があり次第、
先ずグループAの単語を読んでいただきますが、身内や友人に話しかけるようにリラ
ックスした状態で行ってください。グループAを読み終えたら一休みし、次の指示を待
ってからグループBを読んでください。グループBを読み終えたら再度一休みし、指示
があったら文章を読んでください。技術上の問題が生じた場合には、再度読んでいただ
たく場合もあります。

グループA

4) さんすう（算数） -- Sansuu (math)
5) かさ（傘） -- kasa (umbrella)
6) せんせい（先生） -- Sensei (teacher)

グループB

4) すもう（相撲） -- Sumoo (sumo wrestling)
5) みそ（味噌） -- miso (soybean paste)
6) そこです -- soko desu (... is there)

文章

ささきさん、そこにすわってください。 （佐々木さん、そこに座ってください。）

Sasaki-san, soko ni suwatte kudasai -
(Mr. (or Ms.) Sasaki, please sit down there)
APPENDIX C

DEMOGRAPHIC SURVEY

This demographic information will be used to determine if there are correlational characteristics shared by people with specific ways of articulating consonants

Please circle the response or fill in the blank.

Sex:  Male  Female

Age: __________

I was born in: city __________________________ state____________

I attended elementary school in: city ______________ state ______________

List other places you have lived between the time you were 1 and 5 years old:

My first language is: __Japanese
__ other (please specify) __________________________

First language of mother __ Japanese
__ Other (please specify) __________________________

First language of father __ English
__ Other (please specify) __________________________

Languages spoken by mother __________________________

Languages spoken by father __________________________

Languages used in home (as child) __ Japanese
__ other (please specify)

Languages other than Japanese: __________________________
APPENDIX D

DEMOGRAPHIC SURVEY IN JAPANESE

人口統計学的調査

人口統計学的調査は、子音の独特な発音と民族との間に相関的特徴があるかどうかを判断するために使用されます。

回答を丸で囲むか、または空欄に記入してください。

性別：男 女

年齢：18〜19歳 20歳代 30歳代 40歳代 50歳代 60歳代 70歳代

出生地：都道府県 ＿＿＿＿＿＿＿ 市 ＿＿＿＿＿

小学校：都道府県 ＿＿＿＿＿＿＿ 市 ＿＿＿＿＿

上記以外で、1歳から5歳までの間に住んだことのある地名を記入してください。

最初に話した言語：_ 日本語

_ その他（具体的に） ＿＿＿＿＿＿＿

母親が最初に話した言語：_ 日本語

_ その他（具体的に） ＿＿＿＿＿＿＿

父親が最初に話した言語：_ 日本語

_ その他（具体的に） ＿＿＿＿＿＿＿

母親が話した言語： ＿＿＿＿＿＿＿

_ ＿＿＿＿＿＿＿

父親が話した言語： ＿＿＿＿＿＿＿

_ ＿＿＿＿＿＿＿

子供時代に家で使用した言語：_ 日本語

_ その他（具体的に） ＿＿＿＿＿＿＿

日本語以外の言語： ＿＿＿＿＿＿＿

_ ＿＿＿＿＿＿＿
Greg Raver-Lampman  
Email: Glamp003@odu.edu  
Master of Linguistics Program  
Department of English  
5000 Batten Arts and Letters  
Old Dominion University  Norfolk, VA 23529

EDUCATION:  Bachelor of Arts, Comparative Literature, University of California, Berkeley  
Master of Arts, Applied Linguistics (TESOL), May 2012 (pending)

SKILLS:  Teaching English as second or foreign language.

EXPERIENCE: MARKETING AND PUBLIC RELATIONS  
2007-PRESENT: Public Relations Manager, Children’s Hospital of The King’s Daughters

TEACHING:  
2011: English Language Instructor, Old Dominion University Community Course  
2011: English Language Instructor, Old Dominion University Thai Ph.D. student exchange program  
2009: ODU English language tutor  
2006-2008: Adjunct Professor of Journalism and Communications, Christopher Newport University, Newport News, Virginia  
2005-2006: Adjunct Professor of English language for Russian and Ukrainian speakers, University Ukraina, Kiev, Ukraine  
2005-2006: Instructor, American English Center, Kiev, Ukraine

PUBLICATIONS AND PRESENTATIONS:  