Utilizing Audiovisual Stimuli in the Classroom to Facilitate Pronunciation of French Stop Consonants

by Caleb J. Pecue, Bachelor of Arts

A Thesis Submitted in Partial
Fulfillment of the Requirements
for the Degree of
Master of Arts
in the field of TESL

Advisory Committee:

Kristine Hildebrandt, Chair

Seran Aktuna

Larry LaFond

Graduate School
Southern Illinois University Edwardsville
May, 2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

How one acquires a second language has been the topic of discussion for the past century. From vocabulary-acquisition to grammar-acquisition, to acquisition of conversation, many researchers have explored the intricacies of learning a second language. Many of these studies couple production and perception experiments together; however, in this thesis, I only study the production aspect. The perception element of this study should be studied in future experiments. In addition, if production comes after perception, it is unlikely that students perceive French stop consonants as different from English stop consonants.

Many researchers who study SLA recognize that production and perception can proceed at different rates, and in different ways. This study focuses on production; particularly how second language learners of French (English L1) produce voiceless stop consonants in word-initial position.

French stops are unlike English stops in the way they are aspirated. While French voiceless stops /p, t, k/ are realized as unaspirated or unreleased [p̚, t̚, k̚], English voiceless stops are realized as aspirated [p^h, t^h, k^h], specifically, in word-initial position. However, this is not something that can be easily taught given that aspiration is a sub-phonemic and not easily distinguishable.

Because audiovisual stimuli exposure has been shown to be beneficial to, specifically, vocabulary-acquisition, and somewhat mixed reviews to grammar-acquisition, this study expands the idea to target-like production-acquisition of stop consonants. This study takes a longitudinal look at the production of stop consonants in English learners of French and the benefits, if any, to utilizing audiovisual stimuli in the classroom as a facilitator to acquiring

target-like production. Being that this is an expansion on the audiovisual stimuli exposure studies, much of the literature on the specific topic is indirectly related.

Chapter Two of this thesis reviews three types of literature: a) audiovisual stimuli exposure in L2 classrooms, b) voice-onset time for both English and French, and c) second language teaching of pronunciation (CALL). It includes an introduction to voice onset time (VOT), a discussion on place of articulation (POA) as well as the experiments conducted on the benefits to audiovisual stimuli for various types of acquisition: vocabulary and grammar.

Chapters Three through Six focus on the study and the results. Chapter Three presents the methodology and steps took to obtain the data. Chapter Four present the results for the native speakers recorded: French and monolingual English speakers to form a basis of comparison. Chapters Five and Six present the results for the learners. Chapter Five examines first all of the learners as an undistinguished group, then it looks at the two groups, control and experimental, independently, finally it compares the two independent groups to one another. Chapter Six examines the two groups across their proficiency levels: first, at the macro level, looking at the learners as a whole, then at the micro level, looking at the experimental group in depth. Finally, subjects whose productions support the hypothesis are compared with those whose productions do not support the hypothesis.

Chapter Seven concludes this paper by summarizing the findings with a discussion on further studies and the limiting factors of this study.

CHAPTER II

LITERATURE REVIEW

Section 2.1: Introduction

The goal of this thesis is to analyze whether or not L2 target-like production is facilitated by the exposure to audiovisual stimuli, specifically, whether or not learners are able to more accurately produce French-like stop VOT's through time and with exposure to particular instructional stimuli. Throughout this study, prominent questions were asked and examined:

1. Which group shows more progression towards French-like VOT in production across sessions, beginner-levels or intermediate-levels?
2. Is there a noticeable difference between VOT in words produced in isolation versus in carrier phrases?
3. Do we see a continuous, linear progression in French-like VOT through time, or does the change plateau in this progression between Sessions 1 and 4?
4. What are the phonetic differences between stops in L1 French and English?

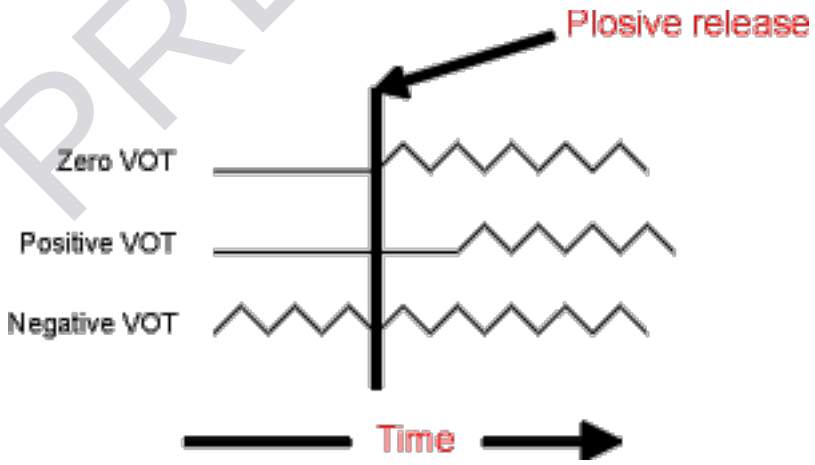
In this chapter, I will present selected review of second language acquisition (SLA) literature through exposure to audiovisual stimuli and on voiceless stop consonant VOT duration differences between English and French. In addition, I will present research on Computer Assisted Language Learning (CALL), which has connections to this study.

Section 2.2: Voiceless Stop Consonants, English and French VOT

One way in which stop consonants may be described is in terms of their voicing and their aspiration. Whalen et al. (2007) define aspiration phonetically as a function of Voice Onset Time (VOT), namely, "the time between the onset of laryngeal vibration and the

release of a stop" (p. 341). From an impressionistic perspective, aspirated voiceless stops are produced with an extra "puff" of air upon the release of the stop closure, while unaspirated voiceless stops lack this extra "puff." VOT is considered positive when the stops are released prior to voice onset and negative when the voicing onset precedes the release. Refer to Figure 2.1 for a descriptive drawing of VOT. The wavy line indicates the voicing, whereas the flat part of the line prior to it (in the case of Zero VOT and Positive VOT) is the period before the voicing, also known as voicing lag. Zero VOT is an example of French VOT, when the plosive is released, the voicing occurs directly after, resulting in not duration of lag or aspiration. English VOT is more associated with Positive VOT, where the plosive is released and aspiration occurs followed by a lag, then voicing. Negative VOT occurs when the vowel voicing occurs before the release of the plosive.

Figure 2.1: Voice Onset Time Depiction



(What is Voice Onset Time?. Retrieved April 20, 2015 from http://www.phon.ucl.ac.uk/home/johnm/siphtra/plostut2/plostut2-2.htm)

French and English stops show an effect of place of articulation, with velars showing longer VOT than bilabials or alveolars, (Lisker and Abramson, 1967). This is also noted by

Ian Maddieson (1997) to be universal in languages that all three places of articulation in their sound inventory, regardless of whether aspiration is phonemic or sub-phonemic (part of allophonic alternations):

> In contrast to this language-specific pattern, the fact that the duration of the aspiration is on average longer with the velar plosive than the bilabial is usually attributed to factors that are inherent in the use of this place contrast (621).

Nearey and Rochet (1994; p.4) detailed the average VOTs for English speakers for the three places of articulation: bilabial, alveolar, and velar. They found that the preceding vowel could affect the duration of voiceless stop consonant's VOT. However, /k/-initial stop consonants were averagely higher than /p/- or /t/-initial, which supports the claim made by Lisker and Abramson (1967) and supports Ian Maddieson's (1997) universal claim. Nearey and Rochet reported the following values as the averages of each POA for English speakers: /p/-initial, 67.4ms, /t/-initial, 73.5ms, and /k/-initial, 79.0ms. For French speakers, the average VOT was as follows: /p/-initial, 31.5ms, /t/-initial, 35.0ms, /k/-initial, 46.3ms.

Unlike English, French voiceless stops are not aspirated. In general, when the time after the initial burst of a stop and before laryngeal vibration is less than 20ms, then it is considered "short-lag," while those longer than 25ms are called "long-lag." In French /p, t, k/ are realized with short voicing lag (Lisker & Abramson, 1964; Caramazza and Yeni-Komshian, 1974). The difference to here is aspiration is the puff of air that occurs when producing a voiceless stop consonant and the lag is the time it takes for the voicing of the vowel to occur. The shorter the lag, the quieter the burst of air will be.

In addition Lisker and Abramson (1964, 1967) have shown that in citation speech, VOT differentiates stops of different phonological voicing classes in initial prestressed position in English and in 10 other languages. Abdelli-Beruh (2004) quotes Lisker and

Abramson (1964) in saying that French and English differ greatly in the manner in which they instantiate the phonological voicing distinction of /p, t, k/. English initial prestressed /p, t, k/ are produced with long voicing lag and are phonetically realized as [p^h, t^h, k^h]. Again, this makes English stop production different from French.

In the Caramazza et al.'s production experiment (1973), VOT between English and French speakers was analyzed. The study contained three groups (Canadian French-speaking monolinguals, Canadian English-speaking monolinguals, and Canadian bilinguals). With the voiceless stops, the monolingual French speakers produced short-lag (<20ms) VOTs, whereas the English monolinguals produced long-lag (>25ms) VOTs. Table 2.1 shows the averages of the different groups across the POAs.

(adapted from Caramazza et al., 1973).

Table 2.1: Average VOT (in ms) of Voiceless Stops between French and English

	Monolingual French	Bilingual French	Bilingual English	Monolingual English
/p/	18	20	39	63
/t/	23	28	48	70
/k/	32	35	67	90

Table 2.1 shows that the bilingual group produced voiceless stops unlike either monolingual group. Recall, Nearey & Rochet's study's averages of /p/-initial, 67.4ms, /t/-initial, 73.5ms, and /k/-initial, 79.0ms for monolingual English speakers. This shows that bilinguals may have an in-between pronunciation, which is neither fully English-like nor fully French-like. This could imply that learning a language at a younger age as a bilingual

may have greater advantages for pronunciation. I will briefly discuss this, the Critical Period Hypothesis, in a future section.

Laeufer (1992, 1996) suggested that the differences between French and English are the realization of /p, t, k/ could be contributed to the fact that English is a stress-based language and French is not. Delattre (1951) stated, "in French, the voicing contrast is carried more by the characteristics of the consonant and, in particular, by the presence/absence of pulsing during the closure" (p. 417). In addition, French stops are fully released, accompanied by audible bursts in utterance-final position (Delattre, 1951; Kohler, 1979); however, English final stops are often partially devoiced (Flege & Brown, 1982) and unreleased (Rositzke, 1943).

The differences in the pronunciation of voiceless stop consonants between French and English may be considered important because when a learner fails to produce the French-like version, their pronunciation becomes marked. A marked pronunciation could be a result of the transfer of L1 sounds into the L2, or even an interlanguage sound (Selinker, 1972). At any rate, it is what contributes to one's foreign accent, which builds one's identity in language learning contexts. At the strong end of the Critical Period Hypothesis spectrum, Lenneberg (1967) states that it would be virtually impossible for adults to acquire native like pronunciation in a foreign language. The Critical Period refers to the ages of around 6-7 years old. This is important to note because if the strong version of the Critical Period Hypothesis were true, then we would not expect to see any student acquiring native-like production in the following study.

In some instances, a foreign accent can be harmful; Derwing (2003) conducted a study on what ESL students say about their accents. Her study took place in Edmonton, Alberta, Canada, which is primarily monolingual. The participants were 100 adult ESL

immigrants who were of varying minorities. This study showed that nearly one-third of the participants stated that they felt they were discriminated against because of their accent. In other cases, this can be even more harmful. Gass (2013) cites the New International Version of The Holy Bible, wherein Chapter 12 of Judges the story of Ephraimites and the Gileadites was told. In this story, in order to detect who was a fleeing Ephraimite and who was not, the Gileadites set up a linguistic test for the Ephraimites, who tradition says could not pronounce the sound /ʃ/ *'sh'*. So, the Gileadites would ask the fleeing men to say the word '*Sibboleth'* and if they could not say it correctly they would be seize them and kill them at the fords of Jordan. It's purported that forty-two thousand Ephraimites were killed at that time (p. 100). This may be an extreme example, however, stereotyping unfortunately can occur still. This can to lead to the learner feeling upset, which could lead to the learner ultimately giving up.

This aforementioned set of literature illustrates the main differences between French and English voiceless stop consonants, and also provides some additional phonetic details about VOT itself in languages where it is distinct. In the next section, I will explore the approaches to second language teaching in addition to studies on multimedia exposure for facilitating L2 acquisition.

Section 2.3: Second Language Teaching Approaches

The field of Second Language Acquisition (SLA) has many different methodologies and positions concerning how and when language learning takes place as well as different methodologies for measuring acquisition. Krashen (1981), for instance, states that acquisition only takes place when input is comprehensible to the learner, as an unconscious process. Pavakanun and d'Ydewalle (1992) infer this to mean that if adults are to acquire a second language in the same way that children do, incidental exposure to another language will lead

them to gradually acquire the language. Vanachter, De Bruycker, and d'Ydewalle (2002) examined whether or not participants who watched a foreign spoken movie would obtain new lexicon by the exposure to the film. They tested their hypothesis by using a sentence recognition test, where sentences to be recognized were directly cut from the soundtrack of the movie. In addition, they mixed in words and sentences that were from other parts of the movie, which the participants did not watch. To see whether or not subtitles had an effect some of the movies where shown with the subtitles and some were shown without them. Their study showed that children seemed to ignore the foreign subtitling when the FL soundtrack was used, Dutch in that case. In d'Ydewalle and Van de Poel (2002), they studied German movies; the adults were shown to perform better on the sentence recognition test than children.

Lonergan (1984) noted that visual media are highly motivating for students and also contextualize language development by exposure to authentic and meaningful models of language use. Although, this study is more of a testament to perception learning, the relation to production can be linked because if one perceives a feature, they may be more likely to produce said feature. In addition, Wood (1996) suggested that it's the narrative element of film that makes them so compelling and that film can provide cross-cultural values and linguistic diversity in an otherwise monolingual situation. Wood suggested that students would often persevere through the difficult language in order to see what happens next.

Chapple and Curtis (2000) looked at film as a means of content under the Content-Based Instruction (CBI). In their study they looked at 31 Cantonese, third-year students, who were taking a General Education course taught in English at the Chinese University of Hong Kong, a bi-lingual university. The course ran for a 13-week semester, twice per week. The first session each week was 45 minutes and the second was 90 minutes. The teaching

pedagogy was highly communicative in approach, with the majority of the class time spent in small-group/whole-group discussion of films and the issues that came from or are explored in them.

Eight films were utilized, seven of which were picked by the teacher and one of which was nominated by the students. The films that were chosen varied in cultural elements and interest, as well as artistic appeal and intellectual challenge. The films came from English-speaking countries as well as those from Europe, Hong Kong, China, and Asian countries. Therefore, not all the films were in English, however, if at all available, the subtitles would be put in English. The class would view the films independently, in a learning center, each week outside of the class, before the discussion.

An English language teacher taught the course, however, the course did not have any specific language goals, as the language was very limitedly taught. The research goals were:

1. Do the students believe their English language skills develop?
2. What are their perceptions of the course and of their own academic development?
3. What other skills and knowledge do they feel they acquire?

In order to answer these questions, the researchers asked to have the students rate how far they thought their English language skills had improved throughout the course, by the following six criteria (425):

1. Confidence in expressing themselves in English
2. Ability to express their ideas when speaking English
3. Ability to express their ideas when writing English
4. English listening skills
5. Knowledge and use of English vocabulary
6. English presentation skills.

No explicit instruction was given, except for minor and incidental corrections to pronunciation. Chapple and Curtis found that afterwards their students' responses indicated that their English language skills had increased in all areas, particularly in speaking and listening skills. The researchers also state the limitations to a questionnaire-type study.

Brinton and Gaskill (1978) studied the effect of listening to TV and radio news on improving EFL students' listening comprehension in an ESL/EFL context, comparing independent experiences of an EFL instructor in Germany and an ESL instructor in the United States. The videotaped broadcasts consisted of BBC's "News of the Week," which is a weekly in-depth broadcast.

The students were then given a two-page handout that consisted of the vocabulary gloss and the comprehension questions. Each word was defined and put into a sentence which related to the context. After the comprehension questions, they were presented with a 12-15 minute edited version of the broadcast. The students were then allowed to recheck their answers and make any additional marks that they had left blank initially. The teacher would then ask individuals for the answers, during which time peer correction was encouraged. The comprehension questions consisted of true/false questions as well as multiple choices. In addition, they wrote a small essay, which was turned into the teacher and handed back the next week with corrections. Brinton and Gaskill note that initially the students felt quite overwhelmed but they persevered and reacted favorably to the technique. "No doubt, the technique provided a welcome relief from the rigidity of the grammar-translation method normally employed in the teaching of English at the school" (407).

In the ESL context, radio broadcasts were used. Three different ESL classes at UCLA Extension were analyzed. Two of the classes had five hours of instruction per day as part of an intensive program (one class was "low-intermediate" and one was "advanced"). In those

classes, the radio broadcasts were used three days a week for thirty to forty-five minutes at a time. The third class was an intermediate class, which met for two and half hours twice a week. The radio broadcasts were used for about thirty minutes over a ten-week period. The broadcasts covered short news segments.

For all classes, a short new report of about one minute in length, which included four or five brief news items, was taped and then transcribed. Because of the initial listening difficulty, the transcript seemed necessary. However, there were omissions in part of the transcript as to challenge the students. The lower intermediate class answered five simple short-answer or true/false questions. Once the transcripts were distributed the broadcast was played three times. Then, they were given time to check over their responses and spelling with a student sitting next to them. The instructor then wrote the answers on the board for the students to check from. The responses were neither collected nor graded. Then, a discussion followed.

As the sessions progressed the transcripts would omit more-and-more words. Although the study had no test to determine the value of using news broadcasts in the classrooms, the following observations were made: First, students were enthusiastic about the broadcasts, and rated them highly in the course evaluation. Second, students asked how they could find the stations on their own radios. Third, students listened to radio and television broadcasts more frequently and understood more than they had prior to the class. Fourth, students would report on further developments of which they had heard or read at home.

At their conclusion, Brinton and Gaskill note that one of the greatest advantages of using news broadcasts in the classroom was vocabulary acquisition. As cited in Blatchford (1973), vocabulary is recycled in the consistent use of the newspaper in the ESL classroom; Brinton and Gaskill note the same thing happening in the broadcasts. “The recycling of

vocabulary takes the pressure off both the student and the instructor in that not everything has to be mastered or taught the first time" (411). Brinton and Gaskill note that it is important to not grade the students on listening performance because, for some, the task is so overwhelming initially that adding grades might prove to be totally demoralizing (412).

Brinton and Gaskill (1978, p. 412) conclude by noting that new broadcasts best fulfill student needs for the following reasons:

1. They are timely and relevant.
2. The recycling of vocabulary is more consistent, particularly in news items which reappear over a period of several week.
3. News items provide the student with a more useful core vocabulary, which enables the student to more readily participate in the type of conversations he is likely to encounter in a social situation.
4. The cultural asides which are a by-product of using news broadcasts provide the student with a broader knowledge of the target culture.

Van Lommel et al. (2006) furthered the concept of audiovisual stimuli exposure's facilitation to not only foreign-vocabulary acquisition but also to foreign-grammar acquisition. Van Lommel et al. state that to master a foreign language, grammar must be acquired beyond just the vocabulary (p. 244). Since previous studies had failed to detect foreign-grammar acquisition, Van Lommel et al. conducted an experiment whereby participants took part in one of the nine following audiovisual exposure conditions: 1) FL soundtrack, subtitled in the FL, 2) FL soundtrack, subtitled in the NL, 3) FL soundtrack, not subtitled, 4) NL soundtrack, subtitled in FL, 5) NL soundtrack, subtitled in NL, 6) NL soundtrack, not subtitled, 7) no Soundtrack, subtitled in FL, 8) no Soundtrack, subtitled in NL, and 9) No Soundtrack, not subtitled.

In the first experiment, Van Lommel et al. investigated whether grammar rules of a foreign language are acquired through watching a subtitled movie. They used reversed subtitling where the soundtrack was in the native language and the subtitles were in the foreign language. The foreign language was Esperanto because it is known for simplicity and small number of rules/irregularities.

Sixty-two (34 females and 28 males) Dutch-speaking sixth-graders from a primary school and forty-seven (32 females and 15 males) Dutch-speaking sixth-graders from a secondary school volunteered. *De Premiejager*, a Dutch spoken twenty-five minute cartoon, was subtitled in Esperanto. This film presented some of the grammar rules. The test consisted of 40 multiple-choice items, eight per grammar rule.

Van Lommel et al. found that the older students performed better than the younger students and presenting the rules beforehand enhanced the performance of the secondary school children considerably more than the primary school children. The experiment showed that the participants performed a lot better on the rules that were presented in the movie than the rules that were not presented in the movie; therefore, a significant interaction between items and the movie appeared.

In the second experiment, Esperanto was used as the soundtrack of the film and the native language was used in the subtitles. 94 sixth-graders from primary schools and 84 sixth-graders from secondary schools participated. The film was called *En Somera Vilao*. They found that no rule acquisition through the movie only and a strong effect of advance rule presentation, particularly among the older children.

Van Lommel et al. concluded by stating that watching the movie did not lead to an incidental acquisition of the rules, and in both experiments the performance improved