UNIVERSIDADE FEDERAL DE SANTA CATARINA

PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA CORRESPONDENTE

THE VOT DIMENSION: A BI-DIRECTIONAL EXPERIMENT WITH ENGLISH AND BRAZILIAN- PORTUGUESE STOPS

GUSTAVO VARGAS COHEN

Dissertação submetida à Universidade Federal de Santa Catarina em cumprimento parcial
dos requisitos para obtenção do grau de

MESTRE EM LETRAS

FLORIANÓPOLIS

Agosto de 2004

To Jacob and Zoé.
My parents, who taught me independence
in place of ordered conformity.

ACKNOWLEDGEMENTS
Much gratitude to:
Barbara O. Baptista
Rosana Koerich
Mailce B. M. Fortkamp
Isabella A. Marques
Preta

ABSTRACT

THE VOT DIMENSION: A BI-DIRECTIONAL EXPERIMENT WITH ENGLISH AND BRAZILIAN- PORTUGUESE STOPS

GUSTAVO VARGAS COHEN

UNIVERSIDADE FEDERAL DE SANTA CATARINA
2004

Supervising Professor: Barbara Oughton Baptista

This cross-sectional study of second language phonetic transfer of adult Brazilian-Portuguese speakers of English attempts to provide a description of the influence the achievement of native-like L2 VOT production exerts on the production of Brazilian-Portuguese unaspirated stop consonants /● /, /● / and /♣/. It also seeks a relationship between production and perception. Its theoretical basis lies in Flege's (1995) Speech Learning Model. The production experiment data was collected with 10 advanced users of English as L2. Samples were collected from each participant using two elicitation protocols: translation and sentence completion. Answers were given in the carrier sentence: " I would say _______". The stimuli consisted of three groups of nine words (all two-syllable words containing stress in the first syllable), each group with one voiceless stop /● /, /● / and /♣/, preceding the back vowels ◘♀◘▪ ◘● ◘ and ◘● ◘ - three tokens for each. Measurements of the VOTs of the speech samples were recorded and digitized in the Acoustics Laboratory of the Universidade Federal de Santa Catarina, using the CSL/Computerized Speech Lab and were analyzed in the software Praat. In the perception experiment listeners heard 120 tokens of randomized ◘↖∗◘ and ◘● ∗/ syllables with VOTs that conformed to English and Brazilian-Portuguese values. They were instructed to

identify the stops in the syllables as English /● / or /✎▭. Results suggest the transfer of L2 VOT values to the native language and an a apparent link between production and perception.

Nº de páginas: 58
Nº de palavras: 14.580

RESUMO

Este estudo de transferência fonética de segunda língua em adultos falantes de português-brasileiro tenta fornecer uma descrição da influência que o sucesso na produção de VOT de segunda língua com valores próximos aos nativos exerce na produção de consoantes plosivas não aspiradas /● /, /● / e /🍁/. Alem disso, busca uma relação entre produção e percepção. Sua base teórica se fundamenta no Speech Learning Model (Modelo de Aprendizagem de Fala) de Flege (1995). Os dados do experimento de produção foram coletados com 10 usuários avançados de língua inglesa. As amostras foram coletadas de cada participante usando dois protocolos de elicitação. As respostas foram dadas dentro da frase: “ I would say ______ (eu diria______)”. O estímulo consistiu de três grupos de nove palavras (palavras de duas sílabas contendo tonicidade na primeira sílaba), cada grupo com uma oclusiva surda /● /, /● / e /🍁/, precedendo as vogais posteriores /♀/, /● / e /● / - três produções para cada. As medidas do VOT das amostras de fala foram gravadas e digitalizadas no Laboratório de Acústica da Universidade Federal de Santa Catarina, usando o CSL/Computerized Speech Lab e foram analisadas no software Praat. No experimento de percepção, os participantes ouviram 120 produções aleatórias de sílabas /✎∗/ e /● ∗/ com valores de VOT condizentes aos do inglês e do português-brasileiro. Eles foram instruídos a identificar as oclusivas nas sílabas como o /● / ou /✎/ do inglês. Os resultados sugerem a transferência de valores de VOT de segunda língua para a língua nativa e um elo aparente entre produção e percepção.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Bi-directional phonetic transfer

A widely held view is that phonetic transfer occurs from the speaker's first language (L1) to the second language (L2). Nevertheless it is known at present that L1 and L2 influence one another in a pervasive manner (Flege, 1998). The Yeni-Komshian et al. (1997) study demonstrated this influence through the conclusion that few of the bilinguals that participated on their investigation were able to pronounce both of their languages without a detectable foreign accent. Just one bilingual out of 240, a woman (with age of arrival of 8.5 years) met the "native-like" criterion in both English and Korean.

Moreover, Sancier and Fowler's (1997) findings demonstrated that the utterances of a Brazilian subject's Portuguese were more foreign accented in terms of VOT of stops after a stay in the U.S. than immediately upon return to the U.S. from Brazil. On the other hand, the subject's English utterances did not suffer significant changes after a stay in the U.S.

Additional research is needed to determine if this bi-directional influence is of equal magnitude and permanence. It is important to consider the bilinguals' history of language acquisition and language use. Theoretically, effects of an L2 on L1 production are more clearly evident in environment in which the L1 is non-dominant, or when the L1 has not been used in the recent past. The current study attempts to shed light on this issue through the investigation of how Brazilian - advanced users of American English (AE) as an L2 -

transfer the aspiration of voiceless stop consonants to their speech in Brazilian Portuguese (BP).

This research is different from past studies in that, first, subjects are not true bilinguals, even though all of them had their first contact with the L2 prior to the onset of puberty. Second, their exposure did not come from native L2 production or from a native L2 environment. Third, the elicitation method used in the data collection procedure was not reading, which avoided L1 sound-spelling interference. Translation and sentence completion were the choices to prevent this L1 decoding interference. The investigation used the Voice-Onset Time (VOT) dimension to acoustically measure and analyze the speech samples collected from the 10 participants. The measure of VOT has been found to be highly effective in separating phonemic categories in languages, even though languages differ in the number of phonological categories and in the other phonetic features assigned to them, as is the case of AE and BP.

This study addresses the following two research questions:

1) Do highly advanced Brazilian-Portuguese speakers of English as second language transfer the aspiration of English voiceless stop consonants to the otherwise unaspirated BP stops?
2) If they do, what is the magnitude and permanence of that influence?
3) Is there a relation between production and perception?

The hypothesis presented in this paper is that those speakers of BP who perform with a production of VOT values that approximate the English norm will transfer these values to the short-lag VOT values of the unaspirated BP voiceless stops.

CHAPTER 2

REVIEW OF LITERATURE

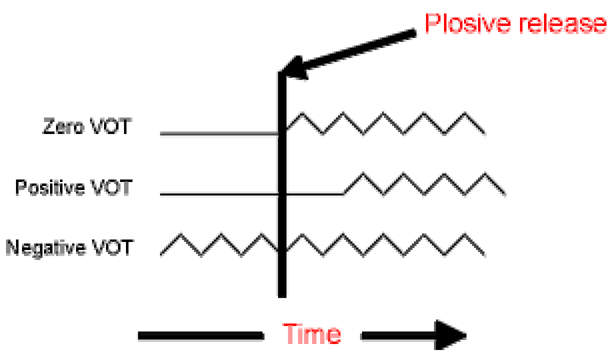
The focus of this chapter is to review empirical research on the variables that affect the acquisition, the production and the perception of the VOT distribution. It also reviews phonetic and phonological theory and the model of language learning on which this study is based.

The chapter is divided into five sections. The first section defines Voice-Onset Time. The second section supplies the theoretical background to this study, with Flege's Speech Learning Model (1995). The third section focuses on the effects of age on VOT production and is divided into two parts. The first part deals with the acquisition of the voicing contrast by children and the second part examines the effects of age on second language VOT production. Section four focuses on the temporal variable of speaking rate and the changes it inflicts on VOT values in perception and production. Finally, section five outlines how several segmental factors, internal and external to syllable structure, influence the VOT environment.

2.1. Definition of VOT

According to Abramson and Lisker (1964) VOT is defined as the time between the release burst of the stop consonant and the onset of periodicity of the following segment, i.e., the vowel. This temporal distance, measured in milliseconds (ms) is referred to in terms of 'lag' and is measured in positive values. If voicing precedes the burst, it is then called 'lead' and is measured in negative values.

Figure 1. Onset of periodicity of the following segment (vocal folds vibration) described by Abramson and Lisker (1964) represented by dented line.



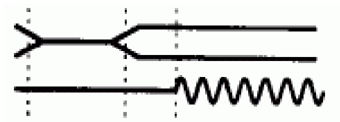
Abramson and Lisker (1964) examined VOT measurements for voicing of word-initial stops produced by speakers of 11 different languages (Dutch, Spanish, Hungarian, Tamil, Cantonese, English, Eastern Armenian, Thai, Korean, Hindi and Marathi) and found that different languages make use of different points along the VOT continuum, as in the following diagram, Figure 2:

Figure 2. Voicing Categories. The top half represents the closing and opening of a stop consonant in the mouth and the bottom half represents the state of the vocal folds; that is, a straight line denotes voicelessness and a wavy line, voicing. Each letter represents one stop voicing category.

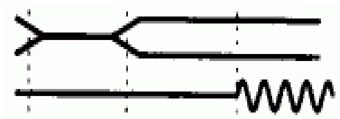
A. Fully voiced (lead)

B. Partially voiced

C. Voiceless unaspirated (short-lag)



D. Aspirated



E. Strongly aspirated (long-lag)

Nevertheless, for ease of representation in linguistic research, stop categories of all languages are said to fall into three general VOT ranges, of which the mean values (pooled across place of articulation) are approximately –100 ms (lead voicing), +10 ms (short-lag) and +75ms (long-lag) (Abramson &Lisker, 1964, p.403).

Thus, different languages carve their voicing categories in different places. From an articulatory point of view, the degree of aspiration, or the amount of ‘lag’, will depend on the degree of glottal aperture during the closure. The greater the opening of the vocal folds during the production of a stop, the longer the amount of following aspiration (Ladefoged, 2001, p. 127).

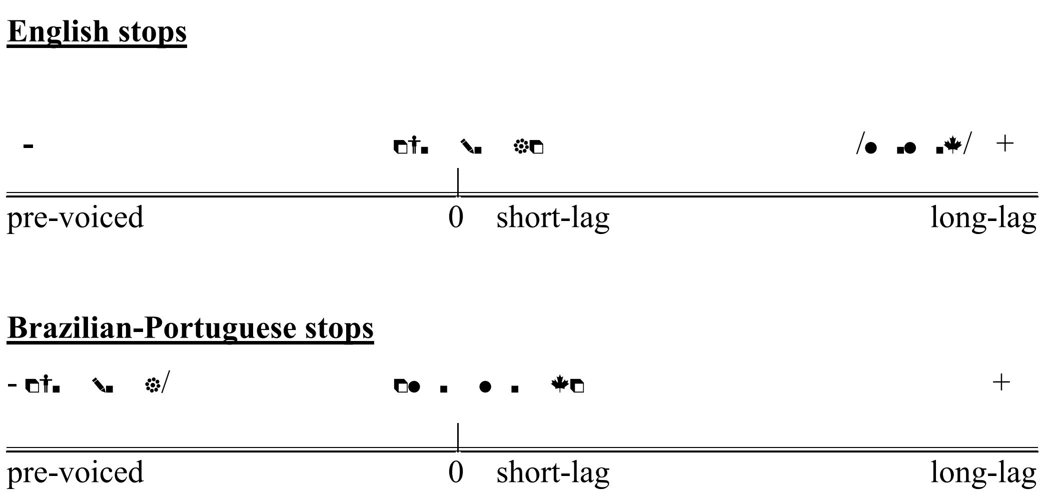
As for the voiced stop consonants, still from an articulatory perspective, production of voicing during /†/ lasts longer than during /✎/ or /❁/ because in the bilabial area there is a fairly large space above the glottis. Air from the lungs can flow through the glottis for a relatively longer period of time before the pressure above the glottis begins to approach that of the air in the lungs. The vocal folds can be kept vibrating through this period. On the other hand, during the production of the velar, there is only a small space above the glottis into which air can flow, which is why voicing can be maintained only briefly. Languages often fail to have fully voiced velar stops (Ladefoged, 2001, p. 130).

Cho and Ladefoged (1999) defined VOT from an articulatory rather than from an acoustic point of view. They adopted this definition due to the active role given to the voluntary initiation of gestures by speakers for the realization of a particular timing for vocal fold vibration. They define it as ‘the time between the *initiation* of the articulatory gesture responsible for the release of a closure and the *initiation* of the laryngeal gesture responsible for vocal fold vibration’ (1999: 225).

While both English and Portuguese fall into the two-category group of languages in terms of the number of voicing categories they contain for stops, they distribute their VOT patterns differently.

As can be noted from the diagram, the VOT range for BP voiceless stops falls nearly within the range of voiced stops in English, while voiced BP stops and voiceless English ones are each at an extreme end of the continuum. As can be seen in Figure 3, stop

Figure 3. Representation of the VOT continuum that shows the relationship between English and Brazilian-Portuguese stops.



consonants are not aspirated in BP. The English word '*two*' and the BP word '*tu*' (personal pronoun *you*) might both be given the broad transcription [tu], but they differ in the presence or absence of aspiration. In many languages, aspiration can change the meaning of a word. For these languages (Thai, for instance) the h-diacritic would need to be included even in broad transcriptions.

One of the most characteristic features of an English accent in BP is the aspiration of stop consonants that should not be aspirated. Similarly, failure to aspirate stops in the appropriate environments can contribute to a Portuguese accent in English. For the sake of representation, if more phonetic detail is required in a transcription, the phoneme /†/ of English can be specified as completely voiceless through [†̥], in, for instance, "that **b**oy". Similarly, the representation of the aspirated /● / that occurs in "pie" is [● ⁻♀✈]. The unaspirated /● / as in "spy" is indicated by [● ● ♀✈] (Ladefoged, 2001, p. 128).

As far as phonetic environment is concerned, in English, voiceless stops are most strikingly aspirated word-initially, before a stressed vowel as in 'pea', 'tea', and 'key'. It is less evident, however, in intervocalic position, as in 'happy', 'natty', or 'lackey' (there is variation across accent). Voiceless stops following /● / as in 'spare', 'stare' and 'scare' are not aspirated. In BP, voiceless stops are not aspirated, regardless of their phonetic environment, and they all fall within the short-lag category (Figure 2).

In short, what appears to be a consistent distinction may be quite variable, and the precise cues that differentiate, for example, 'tie' from 'die' may not be at all the same as those that distinguish 'matter' and 'madder' or 'mat' and 'mad' (Clark & Yallop, 1990, p. 91).

2.2. Acquisition of second language phonetics and phonology

Much of the research that has been conducted up to the present moment on the acquisition of VOT deals with the establishment and maintenance of phonetic categories. Several investigations concerning acquisition of second language VOT deal with the difficulty in mastering timing relations, whether in perception or in production.

The bulk of these cross-linguistic studies have chosen languages that differ in their exploitation of the VOT continuum, such as English and French (Saerens et al. 1989; Hazan and Boulakia, 1993; Kessinger and Blumstein, 1997), English and Spanish (Flege, 1991; Flege et al. 1994; Bohn and Flege, 1993; Flege and Eefting, 1987, 1988), English and Italian (Flege et al., 1995), English and Arabic (Khattab, 2002), English and Japanese (Riney and Takagi, 1999), and English and Brazilian-Portuguese (Sancier and Fowler, 1997; Major, 1987) and the present study.

Since this investigation studies the mutual influence exerted by speakers' L1 and L2, a second language acquisition model that accounts for this interaction is the one that best fits the needs of this research: James Flege's (1995) Speech learning model.

The Speech Learning Model

According to the Speech Learning Model (SLM)(1995), the L1 and L2 interact with one another through the phonetic categories established for position sensitive allophones of vowels and consonants.

The SLM posits that the ability to learn speech remains intact across the life span, but due to changes attributed to age in the state of development of the L1 phonetic categories at the time the L2 learning commences, older learners become less able to perceive and produce L2 sounds accurately, that is, less like a monolingual speaker of the target language. In this view, foreign accents are the indirect consequence of previous phonetic development, not the result of lost or attenuated speech learning abilities.

The success in producing an authentic sounding L2 sound lies in the ability of the speaker to disassociate L2 utterances from his/her L1 repertoire of phonemes and allophones. The level of success varies considerably across speakers. Successful producers of near-native sounding L2 pronunciation are remarkable and rare. This disassociation minimizes the transfer of phonological features from one language to the other, and is necessary due to the fact that the two languages may contain sounds that are apparently the same, but that in fact are produced with differing articulatory movements.

According to the model, it is easier for early bilinguals to eventually establish a phonetic category for target L2 sounds that are auditorily distinct from the closest L1 sounds, whereas it becomes more difficult for late bilinguals, hypothetically, to establish phonetic categories for 'new' L2 sounds that differ substantially from the closest L1 sound, but not for less distant L2 sounds that might be described as 'similar' to L1 sounds.

According to this proposal, the likelihood of category formation varies inversely with age of L2 learning, but directly as a function of perceived cross-language phonetic distance.

Once a new category is established, the exemplar of this category will shift away from its perceptually linked L1 counterpart. The SLM explains that this forced distance in phonetic space occurs as a result of a need for maintaining contrast between the two similar sounds, and "will occur only if a new L2 category is relatively close in phonetic space to a pre-existing L1 category" (Flege, 2002, p. 8).

This assertion is consistent with the notion that bilinguals have separate representations for, for example, /🚹/ and /● / that are comparable to those of monolinguals and that they are capable of activating one language and deactivating the other under appropriate circumstances.

One of the hypotheses of the model is that if instances of an L2 sound continue to be identified as its perceptually linked L1 sound, category formation will be blocked. One of the predictions is that a "merged" category will be developed.

The SLM (Flege, 1995) provides a framework through which to examine changes across life span in the ability to produce and perceive consonants and vowels. This model, which starts with the assumption that one's ability to learn speech does not end or deteriorate after the passing of a critical period, posits that the L1 and L2 influence one another, and that this interaction constrains performance accuracy in both the L1 and the L2. According to the SLM, a number of factors determine whether an L2 learner will discern the phonetic difference between an L2 sound and the closest non-identical sound in the L1. Among these factors are two subject factors: the state of development of the L1 phonetic system at the time L2 learning commences (which is often indexed by age of arrival in an L2-speaking environment) and amount of experience in the L2 (often indexed by length of residence in a predominantly L2-speaking environment).

An illustrative example would be that of a Brazilian-Portuguese speaker, for instance, who produces a short-lag /● / (with VOT of approximately +20 ms) and who is learning American English, and consequently the long-lag /● / (with VOT values of about

+80 ms). Such a speaker might “merge the L1/L2 /• / and produce it with compromise values (around 40 or 50 ms)”. By hypothesis “the properties specified by a merged L1-L2 category can be modeled as a probability-density function reflecting all tokens of the perceptually linked L1 and L2 sound that have been experienced” (Flege, Schirru & MacKay, 2002, p.6) with recent tokens perhaps being given greater weight than tokens encountered in the distant past according to Sancier and Fowler (1997).

In short, the SLM posits that, after having established phonetic categories for L1 sounds, both child and adult listeners are increasingly likely to identify L2 sounds that partially resemble corresponding sounds in the L1 (referred to as similar sounds) as being realizations of an L1 category.

Late learners, however, will persist in identifying similar sounds such as Portuguese and English /• / as being exactly the same, whereas early learners will eventually note the acoustic phonetic differences between them. Theoretically, as a result of this, early but not late learners will establish phonetic categories for similar sounds, and early but not late learners will produce them authentically. There are many studies involved in this issue that corroborate this claim and will be presented in the next section.

2..3. Age and VOT

The age at which the voicing contrast is acquired varies across children and across languages. For example, the age at which English-speaking children acquire a voicing contrast in initial stops may vary from under age 1;10 to 2;8 while Spanish speaking children acquire adult-like phonetic voicing contrast after the age of 4.

Among languages with more than two types of voicing contrast for stops, Srivastava Hindi reports the following sequence: voiceless unaspirated stops are acquired at 1;1, pre-voiced at 1;4, voiceless and voiced aspirates at 2;0 (Khattab, 2002).

Long (1990) concludes from a review of previously published studies that an L2 is usually spoken without accent if learning begins by the age of six years, with variable success between the ages of 6 and 12 years, and with a foreign accent if learning begins after the age the age of 12 years.

With respect to contrast at different places of articulation, the results of Macken and Barton (1989) indicate differences by language (or by voicing type): the English (aspiration) contrast is acquired in the order dental>labial>velar, while the Spanish (true voicing contrast) appears first at the labial place of articulation.

The suggestion that lead voicing normally develops late is supported by acquisition data from French, a language similar to Spanish in its use of lead vs. lag contrast. Allen (1985) found, in data elicited from six French children aged between 1;9 and 2;8, that only three out of a total of 193 word-initial stop consonants had measurable lead voicing.

Late acquisition of lead voicing may be due to production difficulties, even though perceptual factors might also play a role. Deuchar and Clark (1996) argued that the lead/lag contrast is less salient acoustically than the short/long lag contrast, because the latter contrast includes perceptual cues in addition to those of voicing. They discuss the evidence that 'lead' VOT differences are less well discriminated by adults than 'lag' VOT differences; furthermore, "infants aged 6-12 months were more sensitive to VOT differences in the 'lag' region of the VOT continuum than in the 'lead' region" (p.334).

Deuchar and Clark (1996) carried out a case study of a child acquiring English and Spanish in England between the ages of 1;7 and 2;3. Analyses were made from the parents'

Spanish production in terms of ‘lag’ and results indicated that VOT measurements were similar to those of the child at age 2;3. Thus it seems likely that it was the ‘lag’ contrasts in the parents’ speech that were important in the child’s development, rather than the presence of ‘lead’ voicing in the father’s speech or its absence in the mother’s. By studying a bilingual child longitudinally, it was possible to establish the relative order of acquisition of the voicing systems of English and Spanish within the same child, where ‘age’ and ‘stage of cognitive development’ are of course held constant.

A hypothesis generated by the SLM is that the phonetic categories needed to produce and perceive L2 sounds can be added readily until the age of 5-6 years, when the phonetic system begins to stabilize. A sensitive period hypothesis offers no insight into what specific speech learning mechanisms or processes are changed or attenuated as humans mature physiologically and develop cognitively. The SLM posits that the basis for a sensitive period is the increasing frequency of equivalence classification by older children and adults compared to young children. A complete description of the SLM falls outside the scope of this study.

2.3.1. L1 VOT Acquisition

During early L1 acquisition stages the usual development seems to be for all stops to be initially produced in the short-lag range. By 24 months, VOT distinctions usually start to emerge, and the production is extended to the long lag and pre-voiced ranges (Khattab, 2002).

Children are known to produce VOT with longer duration and more variability than adults. Adult-like consistency is usually achieved when reductions in the duration of speech sounds, and in variability, gradually take place as children become older, until approximately 10 to 12 years of age. Still, there are important individual differences in the developmental patterns of children, and a gradual decrease in the duration of sounds is not always the norm (Khattab, 2002).

The age of acquisition of adult-like patterns varies according to the speaker's language. For example, children who acquire a native language that contrasts short-lag with long-lag, such as English, master the adult pattern earlier than children who acquire a language that contrasts voicing-lead with short-lag category (as in BP). The late acquisition of voicing-lead may be due to the complexity of the articulatory gestures involved in its production, mainly the difficulty of coordination of laryngeal control with a particular supra-laryngeal articulatory gesture (Khattab, 2002).

Studies have been conducted also in the area of articulatory disorders, such as Young and Gilbert (1988), who investigated whether VOT distribution would be different in normal speaking children and in children who exhibited velar fronting. Their study aimed at determining how velar fronters produced minimal word-pairs that were perceived as homonymous, and at determining if children were able to differentiate their own productions of the words. Results demonstrated that they were not able to discriminate differences in their own productions. Nevertheless, individual differences in the ability to make that distinction were evident. The normal speakers identified their own productions at an accuracy level of 83% and no child in the normal group made the discrimination at a 100% level. This suggests that the perceptual task was not easy even for the children with normal articulatory/phonological skill. In their results, VOT data failed to show that the velar fronting group was using VOT length to contrast word-pairs. The velar fronters failed consistently to produce longer VOTs for velar stops, although some of the word-pairs showed longer VOT lengths for both /🍁/ and /❁/.

2.3.2. L2 VOT Acquisition

Age is perhaps one of the most investigated issues in cross-linguistic studies on the acquisition of authentic VOT values. The literature generally presents three types of explanation for age-related effects on speech performance (Flege, Schirru & MacKay, 2002).

The first type posits that because of maturational constraints, as the age of first exposure to the L2 increases, the mechanisms responsible for L1 speech acquisition lose their importance. That is why some researchers assign foreign accent to the passing of a critical period for language learning.

The second type claims that “late bilinguals receive less adequate L2 phonetic input than early learners” (Flege, Schirru & MacKay, 2002, p.3), basically due to social reasons, especially in the United States, where the bulk of second language acquisition research is conducted and, at the same time, a great number of immigrant communities are studied. It is true, however, that the L2 might become the bilingual’s dominant language if it is used more than the L1.

A third type of explanation revolves around language interaction. Should it be possible to hear an L2 speaker producing authentic L2 speech without any trace of his or her L1, one might have the idea that this speaker’s L1 and L2 phonetic systems are completely separate. This notion, however, has been discarded by many theorists.

It is conceivable that young children are more likely than adults to arrive at the "new" category solution when exposed to L2 phones not found in L1, due to the fact that they are still in the process of establishing phonetic categories based on the phonetic input they receive in the L1.

According to Flege (1995), early L2 learners have the potential advantage of being exposed to tokens of, for instance, English [• ⁻],[• ⁻] and [✤⁻] at a time when it may be relatively easy to establish new phonetic categories; moreover, children still have many years to refine the phonetic realization rules necessary to output those phonetic categories.

It has been found that Arabic children aged 2-11 years learning both English and Arabic simultaneously realized /• ▪• / in English words with English-like VOT values (82ms). Although the Arabic children's speech production was not compared to that of age-matched native English Children, this suggests that they may have approximated the VOT norm of English more closely than adult L2 learners in previous studies (Khattab, 2002).

Flege and Hillenbrand (1984) found that none of the L2 speakers (adults) who served as participants in their study were able to produce VOT values that resembled authentic native values because they were only exposed to English after the sensitive period.

Results of these studies report that no difference for VOT was found between English monolinguals and adult subjects who learned the L2 as young children.

In contrast, findings reported in many studies claim that adult L2 learners were unable to realize L2 ▭• ▪ • ▪ ✤▭ authentically, and again, this is directly associated to age-related factors, mainly the age at which subjects examined began learning the L2.

Williams (1980) found that 8 to 10 and 14 to 16 year-old native Spanish (Puerto Rican) children produced English /● / with Spanish-like mean VOT values of about 40 and 20 ms respectively. These values were probably shorter than what would be observed for age-matched native English children, suggesting that even child L2 learners may not realize L2 stops authentically.

The hypothesis of the sensitive period is not supported by the findings in Major (1987), in which adult learners of a second language were able to achieve native-like VOT proficiency.

Flege and Eefting (1987a, 1988) present an explanation that might aid in the comprehension of some of those results. They say that possibly the incomplete approximation to the L2 phonetic norms may have been the result of non-authentic L2 input.

That is why all variables should be so meticulously explored. The conflicting results lead to the conclusion that there is a sensitive period that only facilitates native-like attainment, but does not make it impossible after it. Factors such as "motivation, social acceptance and social distance, personality variables, sex, and oral and auditory capacities" (Leather & James, 1996, p. 270) might aid directly or indirectly in the understanding of this matter.

Leather and James (1996) raise the question of whether authentic L2 production is acquired ahead of near-native L2 perception. Their outcome was taken to indicate that production was in advance of perception.

Flege and Schmidt (1995) had 40 native English subjects and 40 native Spanish late bilinguals rate the members of a VOT continuum for goodness as instances of English /● /. A significant correlation was found to exist between the perceptually preferred VOT values (those with the highest goodness ratings) of the native English subjects and their production of VOT; that is, the subjects who tended to produce relatively long values for /● / tended to perceptually prefer stimuli having longer VOT values than did subjects who produced English /● / with shorter VOT values, $r= 0.536$, $p < 0.001$. The native Spanish participants were divided into "proficient" and "non-proficient" subgroups based on their overall degree of foreign accent in English. The correlation between the VOT values produced by the proficient subjects and their perceptually preferred VOT values was significant, $r=-0.004$, $p >0.10$, while for the non-proficient participants it was not. This finding suggests that perception and production align (or become equilibrated) in adult L2 acquisition.

2.4. Speaking rate and VOT

Speech is an extremely complex acoustic signal that varies as a function of a number of different contexts including changes in speaking rate, different talkers, phonetic contexts, emotional states and others. This variability results in an extremely complex mapping relationship between elements in the acoustic signal and the phonetic representations. This complexity assigns to Speaking Rate (SR) a delicate position in the studies of VOT distribution. This is so true that the literature is often inconsistent in defining VOT categories (some studies claim that long-lag is defined around 70 ms or more, others define long-lag in the vicinity of 40 ms)(Clark & Yallop, 1990).

It is also true that individuals show considerable variation in the extent to which they manifest different temporal patterns in the speech in their particular languages. It is natural to say that some individuals talk faster/slower than others. There is also considerable variation among individuals in what is considered the normal SR condition, and there is even more substantial variation in speakers' fast rate of speech. Naturally, temporal variation occurs within and across speakers.

What causes such differences is a rather delicate question. Also adding to the difficulty in determining possible causes of these changes in pattern is the fact that they are not necessarily independent but can, in at least certain instances, interact with one another.

It is common ground among researchers that changes in speech rate affect the range of VOT values associated with each category. Thus, as SR is slowed, VOT is lengthened; conversely, as SR increases, VOT values shorten. At this point, VOT differences at the voiced-voiceless boundary may be reduced, that is, the two phonetic categories are not equally affected by changes in SR. As speakers slow down, they produce not only longer VOTs, but tend to produce a wider range of VOT values. VOT also varies as a consequence of style of speech, whether it is free oral production, the reading of long passages, the production of citation-form minimal pairs or listing.

VOT values of voiceless stops show a greater influence than those of voiced stops. SR has larger effects on the VOT of long lag stop consonants. Kessinger and Blumstein (1997) investigated the effect of SR on languages with three different phonetic categories of voicing, Thai, French and English. VOT was examined in bilabial and alveolar stops produced in CV(C) words in isolation and in context at slow and fast speech rates. Results revealed that the short-lag category did not change substantially as a result of speaking rate in any of the three languages. The long lag category of Thai and English and the pre-voiced category of Thai and French, however, shifted towards short-lag values in the fast rate condition, which resulted in some overlap between voicing categories.

It is commonly agreed that short-lag stops experience minimal variation (Magloire & Green, 1999; Weismer, 1979; Flege et al., 1996), but for the question of why SR rate influences long-lag tokens more than short-lag, the literature presents two possible answers. One is with respect to production: it has been proposed that the way of controlling the onset of voicing differs for the two types of stops. From an articulatory perspective, short-lag stops involve an adduction of the vocal folds prior to the release of the stop, during the preceding closure interval. This ensures that the vocal folds are closed and slackened during consonant release, resulting in relatively short VOT values. The timing constraints are variable for short-lag stops because the command to adduct the vocal folds can be initiated at any time during the closure interval. As long as they are closed and slackened at the time of consonant release, relatively short VOT values will occur.

The other explanation is that the timing is different and more precise for long-lag stops, in which an opening and closing of the vocal folds as a single gesture is initiated sometime prior to the release of the stop. The coordination between the laryngeal gesture and the oral release are timed such that maximum glottal opening occurs either right before or during the release of the stop consonant.

The differences in articulation between short-lag and long-lag stops would suggest qualitative differences in their VOT values in English. However, at a very fast SR there is considerable overlap in the VOT values for these two types of tokens in comparison to their VOT values at normal SR.

This raises a question regarding the degree to which there is a qualitative difference in the timing of these two types of tokens. If long-lag stops can be initiated so as to produce short VOT values at fast rates of speech, why cannot short-lag stops be initiated so as to produce much longer VOT values at very low rate.

An alternative explanation for the asymmetrical impact of SR in the VOT of long-lag and short-lag tokens is related to perceptual constraints. One problem with this proposal is that VOTs of short-lag and long-lag stops do overlap at fast SR, resulting in potential perceptual ambiguity. However, studies of speech perception show that the perceptual system adjusts to changes in SR by shifting the criterion VOT boundary towards shorter VOT values at a fast SR, reducing the amount of potential ambiguity that might otherwise result from such an overlap. Lisker and Abramson (1967) found that Spanish and English, which both use the short-lag category but use it to implement voiceless and voiced stops, respectively, show a similar small range of VOT values across talkers. A substantial increase in the VOT of voiced stops would result in overlap between the voiced and voiceless categories.

According to Kessinger and Blumstein (1997) short-lag is most stable across SR because the size of the category is small, ranging from +1 ms to about 30 ms VOT. They pose that there is little acoustic space in which exemplars of the category may vary.

The question of why voiced stops in English do not show more spread into lead range might be answered through the observation that using voicing lead may require more difficult changes in articulatory timing during the production of the sound.

One way of answering the question regarding the asymmetry in VOT variability across phonetic categories is to examine the effects of short-lag on VOT in languages like Spanish, which do not utilize long-lag VOT values to make a phoneme contrast.

Schmidt and Flege (1996) investigated the impact of SR on the VOT of voiceless stops /● / and /ŧ/ by Spanish and English monolinguals. For the Spanish monolinguals, the VOT of the short-lag (i.e. voiceless) tokens suffered few changes across SR and for some subjects, actually got longer for faster SRs.

Like Spanish, French uses the lead and the short-lag categories to contrast voiced and voiceless stops. As in Spanish, Kessinger and Blumstein (1997) found no effect of SR on the mean VOT of the short-lag tokens, although there was a significant effect of rate on the VOT of the lead tokens: as SR got slower, the mean duration of lead increased. The researchers did find, however, that the range of VOTs for the short-lag tokens in French was greater than the range for English. They speculated that the increased range in French may have been due to the lack of long-lag category. The results of Schmidt and Flege (1996) and Kessinger and Blumstein (1997) provide preliminary evidence in support of a universal restriction on the amount of variability in short-lag tokens.

Nonetheless, the increased range of VOT for the short-lag stops in French raises questions about the possibility of a universal restriction on variability of short-lag tokens. There is the possibility of differences between English and Spanish and French in respect to variation in the SL category being due to talker variability rather than actual language differences.

One way of eliminating talker differences across the two languages would be to use fluent, early bilinguals, which would result in the same talker producing short-lag stops in both language conditions. In this way, each speaker would serve as his/her own control and talker variation across language type would be effectively neutralized.

Schmidt and Flege (1996) reported on the production of /● / and /● / by early and late Spanish-English bilinguals when speaking English but not while speaking Spanish. They found that the early bilinguals produced mean VOT values similar to English monolinguals across the three SRs for English /● / and /● /. The early bilinguals also produced VOT values across changes in SR that were comparable to English monolinguals for the long-lag stops. However they do not provide any cross-language comparison using the same subjects with respect to the short-lag category.

Magloire and Green (1999) investigated the impact of SR on the productions of short-lag stops across Spanish and English by examining both monolinguals' productions of voiced and voiceless stops, as well as bilinguals' production in both Spanish and English. Their study was designed to allow a comparison between the monolinguals and the bilinguals within a particular language, as well as to compare the production of short-lag stops across both languages for different talkers and for the same talkers. In this way, any differences between the two languages due to talker variability could be isolated.

Since there is no phonetic category beyond 'lead', it is expected that slower SR will produce longer duration of lead in Spanish as happens in French (Kessinger & Blumstein, 1997). A related question of interest is whether there is overlap in the VOT distributions of pre-voiced and short-lag stops in Spanish at faster rates of speech.

Kessinger and Blumstein (1997), however, found no evidence of overlapping between short-lag and lead in French, even at fast SR.

Research on changes in SR in production and perception sometimes yields conflicting results. When speakers talk they do not maintain a constant rate of speech, and that is not only a problem for production; this produces a potential problem for perception as well, because many phonetically relevant acoustic properties are themselves temporal in nature, and change as SR changes.

As speakers slow down, such that overall syllable duration becomes longer, the VOT values associated with the stop consonants at syllable onset also become longer. This suggests that if listeners are to use VOT to distinguish between voiced and voiceless stop categories more effectively, they should treat VOT not absolutely, but in relation to syllable duration. Nevertheless, VOT does not depend only on duration of the syllable, but also on its structure. VOT is longer, for example, in the syllable /● ✈/ than in /● ✈● /. That is because there is a compensation in terms of number of segments and total syllable length.

Several rate effects on the production of word-initial stop consonants are mirrored in perception. Perception studies use stimuli that vary inversely VOT and vowel length. Nevertheless, there is no evidence that VOT and vowel length vary inversely in natural speech, that is, that as VOT increases vowel length decreases (Weismer, 1979). It is also known that, at least insofar as syllable boundaries are concerned, a change in syllable-external rate also affects perception.

Considering speech production and perception, there appears to be some conflict between the acoustic data from speech production and the acoustic parameters manipulated in speech perception experiments.

Speech perception experiments exploring the effects of speaking rate on the perception of VOT typically use stimuli with a constant syllable duration. Syllable duration is used to reflect slow and fast SR and each of the two syllable durations is held constant as VOT is varied. Thus, within each syllable duration, VOT and vowel data are varied inversely, such that as VOT is lengthened, vowel length is shortened. In some studies cited in Kessinger & Blumstein (1997) when VOTs are longest, vowel durations are the shortest, and for some stimuli vowel length has been reduced to as little as 5 ms.

However, the data from speech production studies indicate that most of the change in syllable duration as a function of speaking rate is due to change in vowel duration and not change in VOT . If VOT and vowel length do not vary inversely in speech production, then “the perceptual shifts which have been shown to occur when VOT is lengthened as syllable duration is held constant may not relate to changes in speaking rate per se, but may be due to other perceptual effects” (Kessinger & Blumstein, 1997, p. 118). Namely, a lengthened VOT signals speech produced at slow rate, but a shortened vowel signals speech produced at a fast rate. While the increases in VOT prepared for the stimuli of these studies are consistent with slower speech, the concomitant decrease in vowel duration is not, and in fact is compatible with the opposite – namely a faster SR. In order to provide a true test of the effects of speaking rate on perception, it is necessary to utilize stimuli which more closely reflect the changes observed in production, where “VOT and vowel length do not vary inversely within a speaking rate, and where both VOT and vowel length increase as SR slows” (Kessinger&Blumstein, 1998, p. 127).

While it is evident that VOT judgements and perception are affected when the duration of vowel length and VOT vary inversely, these data cannot be considered measures of the effect of speaking rate on the perception of voicing.

Considering the mean VOT values for /● ✈/ and /🚹✈/ as a function of syllable duration, Kessinger & Blumstein (1997) found that as overall syllable duration increased there was a considerable increase in VOT values for /● ✈/, but only a minimal increase in VOT values for /🚹✈◘. That is, for both voiced and voiceless stop consonants, syllable duration increased as SR slowed, but “VOT values for voiced and voiceless stops were affected asymmetrically, thus, SR had a greater effect on VOT values for voiceless stops than for voiced stops,” (p. 119) as mentioned previously.

These data suggested that while increases in VOT contribute to the increased syllable duration at a slower SR (particularly for voiceless stops) they are not solely accountable for this increase, and in fact, they constitute a smaller part of the total duration increase.

More recent studies relate VOT to age, rate-dependent processing and phonetic environment (the latter will be discussed in the next section with more special attention). Flege et al. (1996) hypothesize that early bilinguals are more likely than late bilinguals to establish a phonetic category for English /● /. If so, then the early bilinguals should show a pattern of rate-dependent processing that is similar or identical to that of English monolinguals, whereas the late bilinguals should differ from English monolinguals.

2.5. Phonetic context and VOT

VOT is conditioned by a series of factors, of segmental and extra-segmental nature. These factors exert systematic influence on VOT measurements. Some examples are variation in stress, identity of the preceding and of the following segment, the presence of clusters, speaking rate (as seen in the previous section), and the nature of the production, that is, in isolation or inside a sentence.

The study of vowel influence on VOT has undergone a complete change since its first experiments with Lisker and Abramson in 1967. They concluded that "vowel identity had no systematic influence on VOT", based on a test in which VOT was measured for ❏🕴■ ✎■ ❂■ ● ■ ● ■ 🍁❏ in combination with twelve syllable nuclei (Lisker & Abramson, 1967, p. 15).

In 1975, Klatt conducted a study from which he concluded that there is a "statistically significant difference in VOT values for voiceless stops, depending on whether a high- or mid- vowel followed the stop" (p. 691). To him, the failure of the Lisker and Abramson (1967) study to detect this vowel effect may be related to the limited scope of the authors' data.

Nowadays it is known that acoustic information conveyed by vowels exerts influence on the features of the preceding and of the following segment. VOT is slightly longer before high vowels like ❐✈▪ ● / and slightly shorter before low ones like ❐♀❐■ To support this claim, Chang (1998) explains that "high vowels engender longer VOT because they offer greater impedance to the air escaping from the mouth (thereby delaying the transglottal pressure differential required for voicing). High vowels offer greater impedance to the air because they have smaller constriction apertures. It also seems to be longer for palatalized stops and before non-nasal sonorants, as in *plane*, *tree* and *close*. Furthermore, vowels are shorter before ❐● ▪ ● ▪ ✤❐ than ❐✝▪ ✎▪ ❁❐▪ and they are inversely related to the closure duration of the consonants (Laeufer, 1996)

VOT also varies as a consequence of change in place of articulation. Experiments regarding the interaction between voice and place of articulation have concluded that velar stops yield longer VOT values than alveolar, dental or bilabial respectively. This covariation has been verified in monolingual and in cross-linguistic studies, such as that of Cho and Ladefoged (1999). Aerodynamic, physiological and gestural timing proposals have been put forth to explain the observed pattern of covariation between place of articulation and VOT.

Firstly, the aerodynamic explanations suggest that the smaller cavity behind more posterior constrictions and/or the concomitant larger cavity in front of the constriction delay initiation of vocal folds vibration. Secondly, on the physiological side, the greater mass and contact area involved in a velar constriction may result in releases that are slower than releases from anterior constrictions. A third proposal is that vocal fold opening duration is constant across place of articulation, with the result that aspiration intervals are increased to account for shorter closure intervals for velar stops (Cho & Ladefoged, 1999).

All these studies fall within acoustically based contextual variation, but not all context effects, at least in speech perception, are acoustically based. Some undergo influence of higher order linguistic variables such as lexical status. Interestingly, lexical status seem to inflict greater effects on VOT in perception rather than in production.

Miller and Dexter (1988) experimented with a series of speech syllables that varied in their initial consonant from /ŧ/ to /● /, specified by a change in VOT. The stimuli consisted, at one end of the continuum, of a word of the language, *beef*, whereas the other end consisted of a non-word, *peef*. Listeners tended to identify stimuli with potentially ambiguous phonetic segments in the vicinity of /ŧ/ - /● / boundary as the real word of the language rather than the non-word. In other words, lexical status produced a perceptual shift in category boundary location.

Flege, Frieda, Walley, and Randazza (1998) attempted to determine if organization at the lexical level exerted a significant influence on non-natives' production of stop consonants in an L2. VOT was measured in the production of 60 English words spoken by Native Spanish (NS) and by the subjects in a native English control group. The dependent variable, VOT, was measured acoustically. The lexical factors examined were "subjective familiarity, estimated age of acquisition, imageability, perceived cross-language cognate status and text frequency". This study showed that two well-known phonetic factors (vowel height and number of syllables) affected the VOT values produced in English /● / by NE and NS subjects in the same way. However, the study provided no evidence that any of the five lexical factors that were examined (frequency, familiarity, cognate status, age of acquisition, and imageability) influenced the NS subjects' VOT values. They concluded that additional research conducted in more naturalistic bilingual speaking contexts would be needed before it can be concluded definitively that L2 segmental production is not influenced importantly by lexical factors.

CHAPTER 3. METHOD

3.1. Research questions

The main objective was to investigate (1) whether highly advanced Brazilian-Portuguese speakers of English as second language alter their VOT values in their L1 as a consequence of constant speech production in L2 and consequently transfer long-lag values to otherwise unaspirated BP stop consonants? Sufficient evidence was found to justify research question (2): What is the magnitude and permanence of that influence? Discarding the accommodation period, real transfer was noticed in L1 speech regardless of recent L2 production. Finally, question number (3): Is there a relation between production and perception? Perception data were collected from three of the total number of participants of this study to check whether the best producers of L2 native-like VOT values were also the best perceivers of VOT category distribution across the two languages.

3.2. Production

3.2.1. Participants

The acoustic data reflecting the production of voiceless stops [illegible] were obtained from a group of graduate students of English and Applied Linguistics from the Universidade Federal de Santa Catarina. Most of them are also EFL teachers. Everyone in the group currently resides in Florianopolis, Santa Catarina, although they all come from different cities and from different states, which means a diversity of regional accents. None of them were paid for their participation in this experiment.

Table 1 summarizes information corresponding to each participant, which includes their identification (1), city and state of birth (2), Sex (3) and their age of first contact with English (4) and (5) their use of English in daily life (there are 8 EFL teachers who are also graduate students of the language and 2 advanced users of English).

Table 1. Summary of information on participants

Participant	City/State	sex	Age 1st contact Eng.	English use
1	Santa Cruz do Sul, RS	F	6 years old	EFL Teacher
2	Passo Fundo, RS	F	7 years old	EFL Teacher
3	Campo Mourão, PR	F	13 years old	EFL Teacher
4	Joinville, SC	F	9 years old	Class/study/socially
5	Porto Alegre, RS	M	7 years old	EFL Teacher
6	Farol, PR	M	10 years old	EFL Teacher
7	Maringá, PR	F	9 years old	EFL Teacher
8	Marialva, PR	F	9 years old	EFL Teacher
9	Curitiba, PR	F	7 years old	EFL Teacher
10	Chapecó, SC	F	11 years old	Class/study

There were a total of 10 participants in the production group (2 male and 8 female), all Brazilians who have Portuguese as their dominant language. Their ages range from 23 to

45. According to self-report, they all have normal hearing and speech with no history of speech or hearing disorder. As for their linguistic experience, 6 had spent different amounts of time in English speaking countries. Their age of first contact with English varied from 6 to 13 years old, i.e., before or approximately at puberty. All of them studied English as foreign language at school and in private language schools, and none had formal or informal education in English-speaking countries during childhood. They all use English on a daily basis for different purposes according to self report. There were also three extra participants that constituted the control group of the production experiment (monolingual BP speakers).

3.2.2. Instrumentation

The production test consisted of two parts. The first part used stimuli in English and, the second, in Brazilian-Portuguese.

The English stimuli consisted of 27 words divided into three groups of nine words. All the words had two syllables containing stress in the first syllable.

The first group was formed by words beginning with the voiceless bilabial stop /● /, the second formed by words beginning with the voiceless alveolar stop /● / and the third formed by words beginning with the voiceless velar stop /✤/. In all groups the stop was followed by the vowels ⧠● ⧠▪ ⧠● / and /♀/: three words for each vowel (see Appendix 1).

The Brazilian-Portuguese stimuli also consisted of 27 disyllable words with stress in the first syllable divided into three groups of nine, following the same pattern of the

English stimuli. The Portuguese /● / is said to be more fronted than alveolar, hence it is considered dental (See Appendix 2).

The original choice of vowels to follow the stop consonants in the preparation of the stimuli were ❐✈❐▪ ❐● ❐ and /♀/, representing the three edges of the BP (base-up triangle) vowel chart, however, assuming that in most of the participants' different regional dialects the /● / in the words beginning with /● ✈/ would be pronounced palatalized, the front vowel /✈/ was replaced by mid-back vowel /● /. Stimuli with /● / have been previously used in second language acquisition literature (Bohn and Flege, 1993) for Spanish.

Besides the 54 stimuli words there were 44 more: 22 distracters in English and 22 in BP. The distracters followed the same pattern (two-syllables with stress on the first one) and were placed randomly among the other stimuli (See Appendices 3 and 4).

3.2.3. Preparation of Speech Samples

Recordings were made individually at different times, according to participants' availability. Six participants were recorded in a Sony tape recorder at their homes, yielding 588 speech samples (98 tokens from each of the 6 participants). Three participants were recorded with a Sony digital mini-disc recorder in a quiet room on campus, yielding 294 speech samples. One participant directly onto the computer with a Sony microphone using the Praat Software at my home, yielding 98 tokens. The audio signal of the first group was digitized in the Laboratory of Acoustic Phonetics of the Universidade Federal de Santa Catarina using the CSL/Computerized Speech Lab model 4300B and converted into .nsp files read and analyzed with Praat software. Samples of the remaining two groups used .wav extension and were also read and analyzed with the Dutch software Praat.

3.2.4. Data Collection Procedure

The 10 participants were tested individually in a production test divided into two parts. The first part was entirely carried out in English. Informal conversations prior to the production test were held in the same language, intentionally, to allow the participants' to accommodate their motor mechanisms and brains to the L2 articulatory phonetic environment.

Participants were instructed to (1) translate the words elicited by the researcher's carrier sentence: "How would you say _______ in English?" or (2) complete the sentence the researcher began whenever they judged necessary, and that differed in many ways.

They were instructed to respond to both elicitation protocols simply using the carrier sentence: I would say _______" with a slight pause between the carrier sentence and the answer. An attempt to control speaking rate was made with the instruction to speak very clearly, if possible at a constant rate and loudness level, as in enunciated speech and to repeat any utterance with which they were dissatisfied for any reason.

The 49 English samples (27 words plus 22 distracters) were collected from each participant in approximately 15 minutes. There was no interval between the first and second part.

The second part started immediately after the first, and was held entirely in Portuguese. Participants already knew what to do as a result of prior instruction. The researcher's cue code-switched to: "Como você diria _________ em Português?, and the participants' answered with the carrier sentence: "Eu diria __________". The second part was completed in approximately 15 minutes, which totaled a little less than 30 minutes for each subject for the entire production part of the data collection. It is arguable that this is not the most appropriate data in that it does not reflect natural utterances in the languages. It is, however preferable to ensure unity of style across languages, even at the expense of naturalness.

3.2.5. Data Analysis Procedures

The measurements were taken in soundwave and spectrographic displays with the software Praat. Data consisted of 540 words (27 English + 27 BP x 10 participants). VOT was determined by measuring the time in milliseconds (ms) from the onset of burst to the onset of voicing, regularly at the start of the first glottal pulse. Some tokens were excluded from analysis because of factors such as noise (due to extremely low quality of data from tape recordings) or mispronunciations.

3.3. Perception

The perception experiment was carried out to determine the possible link between perception and production. Choosing from a variety of VOT values in a range of approximately 200 ms from BP pre-voiced to English long-lag category, participants/listeners were instructed to label VOT distribution as English long-lag /● / or short-lag /\/.

3.3.1. Participants

The participants for the perception test were only 3 of the participants from the production test according to availability. None reported history of hearing disorders. None were paid for their participation in the experiment. One had participated previously in speech research procedures. In the presentation of the results, participants/listeners 1, 2 and 3 correspond respectively to participants from the production experiment.

3.3.2. Stimuli preparation:

Four bilinguals, two male and two female, three of them having Brazilian-Portuguese as their native language and one having American English as her native language, chosen simply by availability, spoke the nonsense words /● ∗● ♀/ and /\∗● ♀/ in Portuguese and /● ∗● ■● / and /\∗● ■● / in English six times each word. The first syllables /● ∗/ and /\∗/ were extracted from the four words in the two languages and analyzed. The average VOT of the 96 samples (4Ss x 2AE x 2BP x 6 productions of each word) was calculated to render 16 tokens (the productions considered most representative of a category) from the four speakers. The 16 tokens were multiplied by ten, resulting in 160 tokens of /\∗/ and /● ∗/ in both languages.

3.3.3. Material

The nonsense words were recorded with a digital mini-disc recorder, normalized in the CoolEdit software going through a process of noise reduction to guarantee the audio quality for the perception test. Tokens were edited and analyzed in the speech-synthesizer software Praat.

3.3.4. Procedure

The perception test was carried out simultaneously with the three participants in English. The two English stops /● / and /\/ were tested under the same condition, i. e., the presence of the two BP stops /● / and /\/. The test was divided in two parts. In the first part the perception of English /● / was tested while intercalated with BP /● / and /\/, and in the second part the perception of English /d/ was tested also intercalated with BP /● / and /\/. For the first part, eighty BP tokens (40 /● / and 40 /\/) were presented juxtaposed with 40 English /● / stimuli (yielding 120 randomized tokens). Listeners were instructed to identify the stop. They indicated their response by circling E/● /, BP/● / or BP/\/ on an answer sheet and were instructed toguess if unsure. In the second part they were instructed to repeat the operation, however this time they indicated the stop by circling E/\/, BP/● / or BP/\/ on an answer sheet and were instructed to guess if unsure. They heard the randomized tokens presented over the desktop computer Genius loud-speakers in a quiet room at a comfortable level of volume. There was an interval of 3 seconds between each token. The entire experiment lasted about 45 minutes, including the instructions and informal conversation prior to the experiment in order to answer occasional questions. There was no interval between the two parts of the test. There was no previous training or practice.

CHAPTER 4. RESULTS

4.1. Results and discussion of the Production Experiment

Results of the first part (English) of the production experiment are presented in Table 2 and results from the second part (BP) are presented in Table 3. Averages were calculated on resulting VOT values that represent each word beginning with a different stop.

Tables 2 and 3 indicate the mean VOT values (in milliseconds) for each subject (numbered from 1 to 10). The number found for each stop consonant is an average of nine productions (three with each vowel 🗀● 🗀▪ 🗀● 🗀 and /♀/) in each language. It is possible to observe a tendency for a difference in VOT caused by place of articulation – the lowest for

/● / and the highest for /● / in English. This tendency is not so clear for Portuguese. Vowel influence on VOT was statistically insignificant. Speaking rate was relatively stable.

Figure 5 demonstrates the average VOT for /● / as produced in English and in BP by the 10 participants. A third bar represents the average of six productions of the same words by three extra participants, BP monolinguals, referred to as the Control group. According to self-report, these participants have had little contact with English and do not use it in their daily lives. This inclusion was a forced attempt to compare production of bilinguals' adaptation after using the L2 with monolinguals' speech production. The control group

Table 2. Average VOT values in English for stop consonant for each participant:

Participant	1	2	3	4	5	6	7	8	9	10
Stop										
/● /	0.041	0.057	0.069	0.039	0.058	0.059	0.112	0.114	0.096	0.077
/● /	0.053	0.057	0.077	0.062	0.070	0.073	0.112	0.115	0.102	0.080
/♣/	0.052	0.052	0.078	0.050	0.062	0.061	0.097	0.113	0.078	0.076

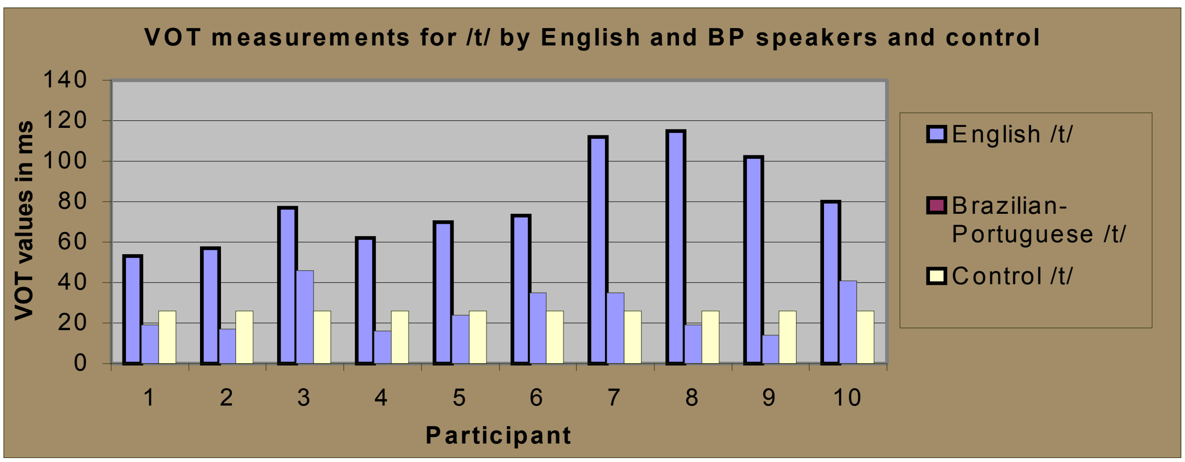
Table 3. Average of VOT values in Brazilian-Portuguese for stop consonant for each participant:

Participant	1	2	3	4	5	6	7	8	9	10
Stop										
/● /	0.020	0.017	0.027	0.019	0.022	0.025	0.025	0.019	0.025	0.025
/● /	0.019	0.017	0.046	0.016	0.024	0.035	0.035	0.019	0.014	0.041

/♣/	0.025	0.023	0.037	0.022	0.027	0.029	0.060	0.054	0.050	0.053

average results (of 6 repetitions of the words beginning with each stop consonant) are represented equally along the graphs. Figures 6 and 7 show the same VOT measurements for /● / and /♣/ respectively.

The dark bar graphs in the middle represent the effect of recent language production. Part of the accommodation was noticeable in the VOT values of the initial



tokens of the BP section. However, since this second part of the test lasted longer than five minutes, the motoric accommodation period was over and only those participants who really transfer are the ones who kept longer VOT values throughout almost the entire second part.

Figure 4. Average VOT production of /● / by bilingual participants and control monolingual group.

The context in which a speech sound is heard clearly inflicts significant changes in the way listeners produce that sound the second time. Jamieson and Cheesman (1987) have worked on the perceptuo-motor adaptation effect. They conducted three experiments:

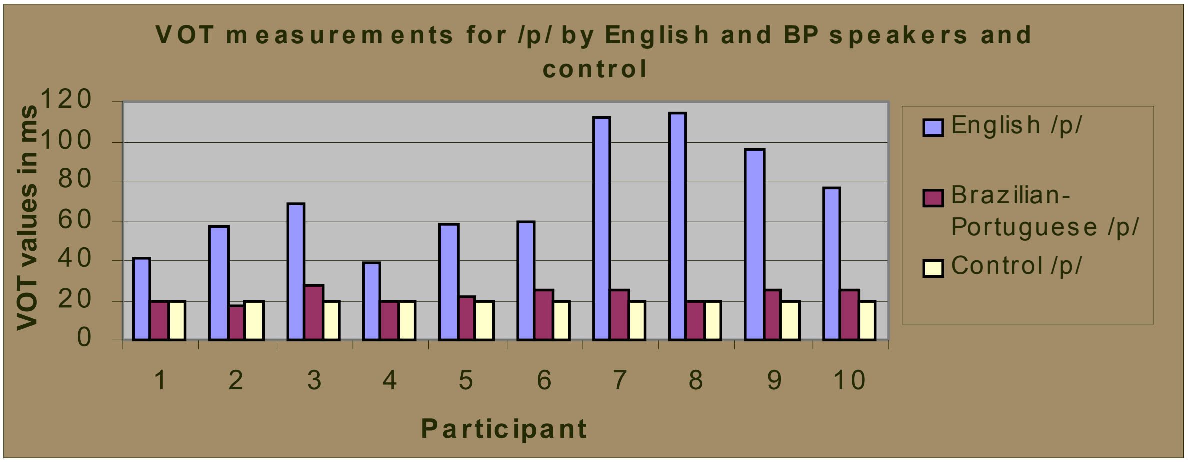


Figure 5. Average VOT production of /● / by bilingual participants and control monolingual group.

Experiment 1 examined VOT for productions of [● ⁻♀], after listening to ✓● ⁻♀☑▪ ✓†♀], or [♀]. Experiment 2 examined VOT for productions of [†♀], after listening to ✓● ⁻♀☑▪ ✓†♀☑▪ or [♀]. Experiment 3 examined the time course of recovery from perceptuo-

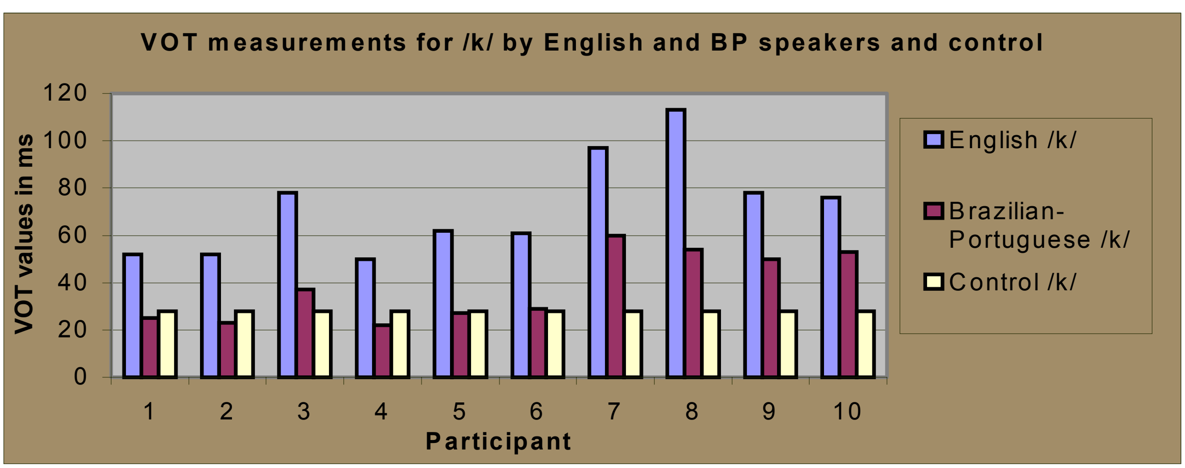


Figure 6. Average VOT production of /♣/ by bilingual participants and control monolingual group.

motor adaptation for [● ¯♀] utterances (p. 17). Results showed that listening to repeated [● ¯♀] sounds made participants shorten their VOT values in the production of the same syllable immediately after having heard it by an average of 7 ms. The adaptation effect was readily inducible and

highly reproducible. In contrast, listening to [ba] sounds had no effect on the production of [● ¯♀] syllables. The production of [ba] after listening to [ba] or [● ¯♀] underwent no changes.

An especially important matter for this study is the fact that the adaptation is short lived, since VOT values for [● ¯♀] produced 30 seconds after adaptation are statistically indistinct from those uttered without adaptation (p. 15). Data collection of the production part took more than 10 minutes, which means that motoric accommodation is discarded; that is, those participants whose L2 influenced their L1 had passed the point of accommodation when speech samples were recorded.

Support for the hypothesis that those speakers of BP who produce VOT values that approximate the English norm will transfer these values to the short-lag VOT values of the unaspirated BP voiceless stops was found in participants 3, 7, 8, 9 and 10, especially for the /♣/ sound. Participants 3, 6, 7 and 10 showed transfer mainly in /● / and participant 3 in /● / to a lesser extent. Participants 1, 2, 4, 5 and 6 produced VOT values for voiceless stops ▫● ▪ ● ▪ ♣▫ in BP that do not show a distinctive difference between what is normally expected of the short-lag categories nor between the monolingual VOT values, which served as control. Most of them have approximated their L1 values to what is considered intermediate for interlanguage, that is, VOTs of approximately 50 ms (between short- and long-lag categories). One could argue that the age factor was an intervening variable. Participants learned their English during childhood but supposedly there was no real contact with native speakers. In addition, native input might have come only unidirectionally, from television and music or several other sources, rather than bidirectionally, through conversation. There are, however, other variables that might have interfered, such as fossilized errors or lack of interest in pronunciation accuracy.

Participants 7, 8 and 9 showed authentic long-lag values in their production but little transfer for /● / and /● /, probably due to the fact that, as graduate students of English, they realized that they were being tested for aspiration, as reported after the recordings, despite the distracters, and that probably had an influence which rendered a somewhat artificial realization of the BP tokens.

Concerning the relation between production and perception, the data collected in this study suggest a stronger link between the two abilities only for participant 3. Schmidt & Flege's (1996) study seeking similar correlation showed the mean goodness ratings that were obtained from four (of ten) Spanish late bilinguals who produced English /● / with Spanish-like short-lag VOT values ranging from 13-18 msec. It also showed that these subjects showed little, if any, effect of the speaking rate manipulation when rating the VOT stimuli for goodness as instances of English /● /. According to their results four subjects who produced English /● / with VOT values ranging from 41-68 msec did show evidence of rate-dependent processing. The VOT values that these four subjects produced in English closely approximated the values obtained for the native English control subjects (range 37-57).

Comparing Schmidt and Flege's (1996) study with the current study, it is possible to conclude that the four subjects of their experiment who were able to produce English /● / accurately and participant 3 of this study have, according to the SLM, established a phonetic category for English /● /. This conclusion, nevertheless, must be viewed in both cases as tentative and subjected to additional testing.

4.2. Results and discussion of the Perception Experiment

There were several hypotheses underlying this experiment. The first was that in a phonetic context that includes /● / tokens of both languages the short-lag Portuguese /● / tokens would somehow sound less voiceless. According to Bohn and Flege (1993), differences between short-lag vs. long-lag stops may be more salient to listeners than between stops with lead and short-lag VOT values because the former contains a wider range of acoustic cues. They suggest that the short vs. long-lag distinction may be more robust psycho-acoustically than adistinction based on lead vs. short-lag. Previous studies have shown that bilingual native speakers of languages in which /● / is realized with short-lag VOT values identify stops with short-lag VOT as voiceless more often in an L1 perceptual set than in an L2 (English) set (Bohn & Flege, 1993).

The second hypothesis was that short-lag /● / would be identified as "● " more often when presented along with short-lag English /\/ tokens than when presented along with long-lag English /● / tokens. If so, the BP /● /s should be labeled "● " less often when juxtaposed to the English /● /s and more often when juxtaposed to the English /\/s.

The third hypothesis was that there would be a relationship between perception and production, in which the participants with better performance in producing native-like stops would also have a more accurate perception of these segments.

Abramson and Lisker (1973) reported that native Spanish listeners may have a secondary discrimination peak in the positive VOT region. This implies sensitivity to a short-lag vs. long-lag contrast in the absence of language-specific input. The reviewed literature does not give information on BP native speakers having more peaks of discrimination, as Spanish speakers do; however, research suggests thatBP speakers are able to discriminate foreign accent, within the VOT dimension, successfully (Sancier & Fowler, 1997, p. 426). That is, besides the inherent lead and short-lag contrast, Brazilians acknowledge the presence of a long-lag category.

Data obtained from the three listeners' responses provided comparison values for the analysis. Results from the first part of the experiment, the English /● / labeling test, are presented in Table 4 with raw numbers and percentages and statistically in Figure 7, followed by discussion. The results from the second part of the experiment, the English /\/

Table 4. Results of the English /● / labeling test. Percentages in parentheses. The phonemes listed at the left side of the table correspond to the tokens heard. Those at the top are the identities attributed to the tokens heard.

Participant		1			2			3		
Seg.	N	E/t/	BP/t/	BP /d/	E/t/	BP/t/	BP /d/	E/t/	BP/t/	BP /d/
E/● ̃▫	40	39	1	0	38	2	0	39	1	0
		(97.5)	(2.5)	(0.0)	(95.0)	(5.0)	(0.0)	(97.5)	(2.5)	(0.0)
BP̃/	40	1	39	0	2	38	0	1	39	0
● /		(2.5)	(97.5)	(0.0)	(5.0)	(95.0)	(0.0)	(2.5)	(97.5)	(0.0)
BP/␀/	40	0	0	40	0	0	40	0	0	40
		(0.0)	(0.0)	(100.0)	(0.0)	(0.0)	(100.0)	(0.0)	(0.0)	(100.0)

test, are presented in Table 5 with raw numbers and percentages and statistically in Figure 8, followed by discussion.

In the English /● / test, there was not much variation among the three participants. Participant 1 identified E/● / correctly 39 times and as BP /● / only once. He identified BP /● / as E/● / once. Participant 2 identified E/● / correctly 38times and as BP /● / twice. He identified BP /● / as E/● / twice. Participant 3 identified E/● / correctly 39 times and as BP /● / only once. He identified BP /● / as E/● / only once.

In the presence of the long-lag /● /, the short-lag ̃/● / was expected to sound more voiced, as mentioned in the first hypothesis; nonetheless, listeners not even once mistook the short-lag /● / for /␀/. They were actually quite consistent in terms of how frequently they labeled the E/● / as voiceless, which apparently corroborates Abramson and Lisker's (1973) findings about a second discrimination peak in the positive region. In addition, as graduate students of English at the time of data collection, the participants had some formal knowledge of phonetic issues, including aspiration. Interestingly, all of them mistook the first token heard, i.e., BP/● /, for E/● / (participant 2 did that twice), possibly because they did not have a point of reference for VOT yet. This study differed from Bohn and Flege (1993) in several ways. Their focus was on Spanish /• / (p.275) and they had four groups of listeners (Spanish and English monolinguals and early and late bilinguals), whereas this study dealt only with native speakers of BP who have massive daily contact with English, referred to here as bilinguals, and focused on English /•˜ / and /\/. In the English /\/ test there was more variation in the results.

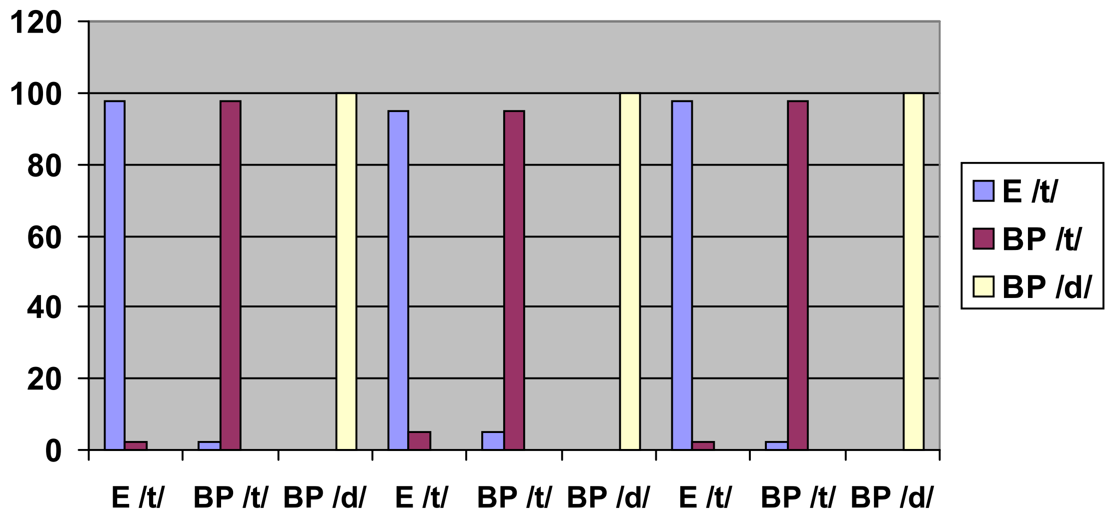


Figure 7. Results of the English /• / labeling test. Identification of each segment listed at the bottom of the figure by participants 1, 2, and 3 from left to right.

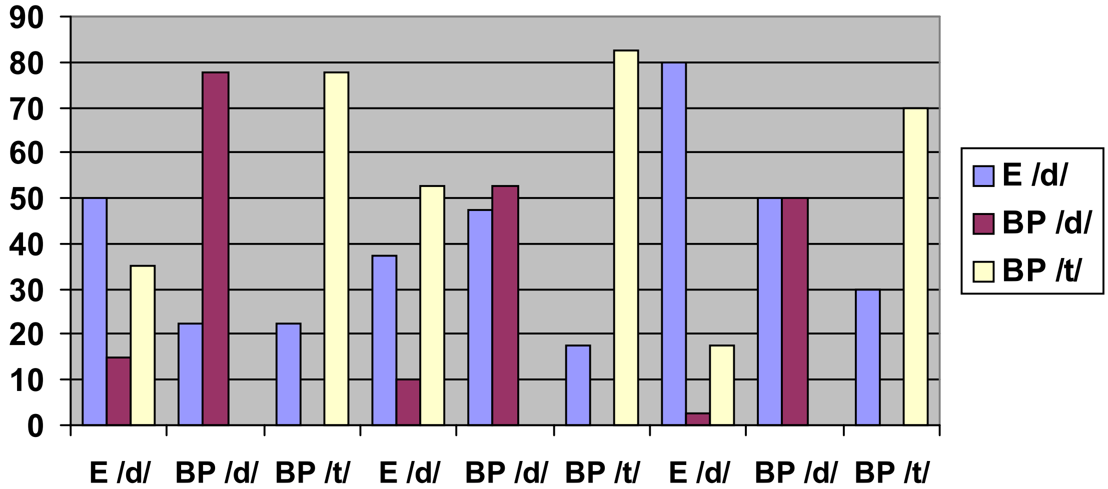


Figure 8. Results of the English /ð/ labeling test. Identification of each segment listed at the bottom of the figure by participants 1, 2, and 3 from left to right.

Participant 1 identified E/ð/ correctly 20 times, as BP/ð/ 6 times, and as BP /● / 14 times. He identified BP/ð/ and BP/● / as E/ð/ 9 times each. Participant 2 identified E/ð/ correctly 15 times, as BP /ð/ 4 times, and as BP /● / 21 times. He identified BP/ð/ as E/ð/ 19 times and BP/● / as E/ð/ 7 times. Participant 3 identified E/ð/ correctly 32 times, as BP /ð/ once, and as BP/● / 7 times. He identified BP/ð/ as E/ð/ 20 times and BP/● / as E/ð/ 12 times. There was a high degree of variability in these results. As mentioned in the second hypothesis, in the presence of E/ð/ the BP/● / was expected to sound more like "● "; however,

Table 5. Results of the English /ð̃/ labeling test. Percentages in parentheses. The phonemes listed at the left side of the table correspond to the tokens heard. Those at the top are the identities attributed to the tokens heard.

Participant		1			2			3		
Seg.	N	E/ð/	BP/ð/	BP	E/ð/	BP/ð/	BP	E/ð/	BP/ð/	BP /● /

				/● /			/● /			
E/✓/	40	20	6	14	15	4	21	32	1	7
		(50.0)	(15.0)	(35.0)	(37.5)	(10.0)	(52.5)	(80.0)	(2.5)	(17.5)
BP/✓/	40	9	31	0	19	21	0	20	20	0
		(22.5)	(77.5)	(0.0)	(47.5)	(52.5)	(0.0)	(50.0)	(50.0)	(0.0)
BP/	40	9	0	31	7	0	33	12	0	28
● /		(22.5)	(0.0)	(77.5)	(17.5)	(0.0)	(82.5)	(30.0)	(0.0)	(70.0)

it did not. BP/● / was identified correctly fewer times in the presence of E/✓/ than in the presence of E/● /.

As for the idea that perception would be related to production, participant 3 had the most native-like production performance, as well as identifying English /✓/ the greatest number of times in the second perception test. Nevertheless, this is very little data for a generalization. In addition, this participant identified BP/✓/ as E/✓/ 20 times and BP/● / as E/✓/ 12 times. Thus, the results provide only weak support for a relationship between perception and production.

5. CONCLUSION

5.1 Significance of findings

While most research in phonetics and phonology of the kind that experiment with languages in contact study the effects of the L1 on the L2, this study gives credit to the little studied area of the opposite influence and it is a contribution to second language acquisition research in that sense. In its attempt to clarify if advanced BP speakers of English as second language transfer the aspiration of English voiceless stops to the otherwise unaspirated BP stops, the results show that 5 out of the 10 participants did transfer in their production performance. This is in fact a very positive result, since technically only those who have really achieved a native like performance showed L2-L1 influence. Support for the hypothesis was found especially for the stop consonant /🍁/. These findings confirm Sancier and Fowler (1997) hypothesis that BP speakers are able to achieve native like L2 pronunciation at least in the dimension of VOT values.

As for the permanence of the transfer, evidence was found in the tokens spoken in the final part of the data collection section, i.e., after the accommodation period had theoretically decreased.

It was also stated that there is a positive relationship between the degree of L2 VOT proficiency and perception, meaning that those who transfer L2 values to their native language do it because they have achieved a high level of L2 use that positively influences their perception.

5.2. Strengths and limitations

The limitations of this study are many. In particular, insufficient data prevented a thorough generalization stemming from the results. The acknowledgment of these limitations has, however, provided contributions to future research. One of them is that research would be more realistic if it examined stop consonants produced spontaneously during a conversation, rather than via test elicitation, which would, nevertheless, involve too many other variables. The influence of English on BP can be noticed better in informal settings, as it was noticed during the conversations after and prior to the data elicitation period for the production part of the experiment.

Furthermore, spectrographic and waveform analysis require sensitive recording equipment. The use of digital quality recording equipment facilitates analysis immensely.

The ideal situation is, naturally, to make all recordings using the same equipment following the same procedures. In the case of the present study, this was not possible due to the fact that the data were collected at different times with different availability of equipment.

One of the disadvantages of cross-sectional studies is the fact that the results are valid for the time of the research only, i.e., students can always grow more accurate and train. A longitudinal study faciliates the task of the researcher while identifying the interlanguage stage of the student and allows him or her to work from there on, to even analyse the effects of training, which is another promising avenue for research to be conducted in this area.

On the other hand, the strengths of the study include the use of state-of-the-art technology to collect and analyse the majority of the data, though not all of it. The review of the literature brings together data that are not found, to my knowledge, in any other place organized in this manner.

Another strong point was the choice of translation and sentence completion rather than 'reading' as elicitation protocol. The latter might lead readers to use their L1 sound-spelling representation rules to decode L2 sounds, and that might cause undesirable L1 interference. Another reason is that reading from lists might lead to a typical 'list' intonation pattern that can have an effect on VOT. Two additional words, that would not be analyzed, would have to be inserted before and after the desired token so as to avoid possible 'end-of-list' effect. Besides, translation as form of elicitation has shown successful results in the literature (Sancier & Fowler, 1997).

Its results contribute to a better understanding of an area that is scientifically complex to study due to the fact that it is so human, and therefore, so full of individual differences.

5.3. Future research

Many questions raised during the study remain unanswered. For example, is there a threshold in production that VOT values must reach in order for L2 transfer to begin to occur? There, one has subject for an entire dissertation; How long does this transfer last? Despite the fact that the time of data collection exceeded the accommodation period, it would be quite sensible to replicate the BP part of the test in order to see how farther the influence lasts under these circumstances. What are the reasons for variation in transfer duration? A promising avenue for production/perception comparison is the search for the effects of training.

REFERENCES

Abramson, A. S. & Lisker (1973) Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics* **1**, 1-8.

Abramson, A. S. & Lisker (1964) Cross-language study of voicing in initial stops: Acoustical measurements, *Word*, **20**, 384-422.

Allen, G.D. (1985). How the young French child avoids the prevoicing problem for word-initial voiced stops, *Journal of Child Language*, **12**, 37-46,

Beddor, P. S. & Gottfried, T. L. (1995) Methodological issues in cross-language speech perception research with adults. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research)pp. 207-232). Timonium, MD: York.

Benkí, J. R. (2001). Place of articulation and first formant transition pattern both affect perception of voicing in English. *Journal of Phonetics*, **29**, 1-22.

Bohn, O.S. & Flege, J. E. (1993). Perceptual switching in Spanish/English bilinguals. *Journal of Phonetics*. **21**, 267-290.

Carney, A. E. et al. (1988). The effect of adventitious deafness on the perception and production of Voice onset time in Thai: a case study. *Language and Speech*, **31:3**, 273-282.

Chang, S. S. Vowel dependent VOT variation, unpublished paper by a group experiment in John Ohala's spring seminar 1998 on Methods on Experimental Phonology at the University of California, Berkeley. (Accessed on August, 2003)
http://trill.berkeley.edu/ICPhS/frameless/errata/steve_chang.pdf

Cho, T. & Ladefoged P. (1999). Variations and universals in VOT: evidence from 18 languages. *Journal of Phonetics.* **27**, 207-229.

Clark, J. & Yallop, C. (1990). *An introduction to phonetics and phonology*. Basil Blackwell, Oxford, UK.

Cooper, W. E. & Nager, R.M. (1975) Perceptuo-motor adaptation to speech: An analysis of bysillabic utterances and a neural model. *Journal of the Acoustical Society of America*, **58**, 256-365.

Davis, K. (1994). Stop voicing in Hindi. *Journal of Phonetics*. **22**, 177-193.

Deuchar, M. & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics*. **20**, 331-345.

Flege, J. E. (1980) Phonetic approximation in second language acquisition. *Language Learning*. **30:1**, 117-134.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*. **15**, 47-65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*. **89:1**, 395-411.

Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In: Strange, W. (Editor) Speech perception and linguistic experience – issues in cross language research, York press, Baltimore, pp. 233-269.

Flege, J. E. (1998). The phonetic study of bilingualism. *Ilha do Desterro*. **35**, 17-26.

Flege, J. E., Frieda, E. M., Walley, A.C., & Randazza, L. A. (1998). Lexical factors and segmental accuracy in second language speech production. *Studies on Second Language Acquisition*, **20**, 155-187.

Flege, J. E. & Eefting, W. (1987a). Production and perception of English stops by native speakers of Spanish. *Journal of Phonetics*. **15**, 67-83.

Flege, J. E. & Eefting, W. (1987b). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, **6**, 185-202.

Flege, J. E. & Eefting, W. (1988). Imitation of a VOT continuum by native speakers of English and Spanish: evidence for phonetic category formation. *Journal of the Acoustical Society of America*. **83:2**, 729-740.

Flege, J. E. & Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language production, *Journal of the Acoustical Society of America*. **76:3**, 708-721.

Flege, J. E., Munro, M. J. & MacKay, I. R. A. (1995). Effects of age of second language learning on the production of English consonants. *Speech Communication*, **16**, 1-26.

Flege, J. E. & Schmidt, A. M. (1995). Native speakers of Spanish show rate-dependent processing of English stop consonants. *Phonetica*. **52**, 90-111.

Flege, J. E., Schmidt, A. M. & Wharton, G. (1996). Age of learning affects rate-dependent processing of stops in second language. *Phonetica*. **53**, 143-161.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40, pp. 467-491.

Hazan, V. L. & Boulakia, G. (1993). Perception and production of a voicing contrast by French-English bilinguals. *Language and Speech*, **36:1**, 17-38.

Jamieson, D. G. & Cheesman, M. F. (1987). The adaptation of produced voice-onset time. *Journal of Phonetics*, **15**, 15-27.

Jäncke, L. (1994) Variability and duration of voice onset time and phonation in stuttering and nonstuttering adults. *Journal of Fluency Disorders,* **19***,* 21-37.

Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*. **60:2**, 286-319.

Kessinger, R. H. & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*. **25**, 143-168.

Kessinger, R. H. & Blumstein, S. E. (1998). Effects of speaking rate on voice-onset time and vowel production: some implications for perception studies. *Journal of Phonetics*. **26**, 117-128.

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children

Kingston, J. & Diehl, R. L. (1994). Phonetic knowledge. *Language*, **70**, 419-454.

Klatt, D. H. (1975). Voice onset time, frication and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research*, **18**, 686-706.

Kuehn, D. P. & Moon, J. B. (1998). Velopharyngeal closure force and levator veli palatini activation levels in varying phonetic contexts. *Journal of Speech, Language, and Hearing Research*, **41**, 51-*62*.

Kuijpers, C. T. L. (1993). Temporal aspects of the voiced-voiceless distinction in speech development of young Dutch children. *Journal of Phonetics*. **21**, 313-327.

Kuijpers, C. T. L. (1996). Perception of the voicing contrast by Dutch children and adults. *Journal of Phonetics*. **24**, 366-383.

Ladefoged, P. (2001). *A course in phonetics*. Heinle & Heinle, Boston, MA.

Laeufer, C. (1996). The acquisition of a complex phonological contrast: voice timing patterns o English initial stops by native French speakers. *Phonetica*. **53**, 86-110.

Laver, J. (1994). *Principles in Phonetics*, Cambridge University Press: Cambridge.

Leather, J. & James, A. (1996). Second language Speech. In Tehie W.C.R. & Bathia T.K. (Eds.) *Handbook of Second Language Acquisition*. San Diego: Academic Press.

Lisker, L. & Abramson, A. (1967). Some effects of context of voice onset time. *Language and Speech*, **10**, 1-28.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, **12**, 251-285.

Macken, M. A. & Barton, C. A. (1989). Phonological universals in language acquisition. 3-22.

Magloire, J. & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica*, **56**, 158-185.

Major, R. C. (1987). English voiceless stop production by speakers of Brazilian Portuguese. *Journal of Phonetics*. **15**, 197-202.

Miller, J. L. (1994). On the internal structure of phonetic categories: a progress report. *Cognition*, **50**, 271-285.

Miller, J. L. & Dexter, E. R. (1988). Effects of speaking rate and lexical status on phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, **14**, 369-378.

Riney, T. J. & Takagi N. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*. **49**, 275-302.

Rochet, B. L. (1995) Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research (pp. 379-410). Timonium. MD: York Press.

Saerens, M., Serniclaes, W. & Beeckmans, R. (1989). Acoustic versus contextual factors in stop voicing perception in spontaneous French. *Language and Speech*, **32:4**, 291-314.

Sancier, M. I. & Fowler, C. A. (1997) Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, **25**, 421-436.

Savithri, S. R. & Rohini, H. Voicing perception in patients with cerebellar pathologies. Department of Speech-Language Sciences, All India Institute of Speech and Hearing, Manasagangothri, Mysore

Schmidt, A.M. & Flege J. E. (1996). Speaking rate effects on stops produced by Spanish and English monolinguals and Spanish/English bilinguals. *Phonetica*, **53**, 162-179.

Smith, B. L. (2002) Effects of Speaking Rate on Temporal Patterns of English**.** *Phonetica* **59**, 232–244.

Tailtelbaum, R., Kishon-Rabin, L. & Hildesheimer, M. (XXXX). Acoustic cues to the perception of initial voicing in Hebrew. Department of Communication Disorders, Sackler School of Medicine, Tel Aviv University

Weismer, G. (1979). Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics*. **7**, 197-204.

Williams, L. (1980). Phonetic Variation as a function of second-language learning. In Yeni-Komshianm G. H., Kavanagh, J. F., & Ferguson, C. A. (Eds.), *Child Phonology*, Vol. **2**: Perception (pp. 185-215). New York: Academic Press.

Yeni-Komshian, G., Flege, J.E., and Liu, H. (1997). Pronunciation proficiency in L1 and L2 among Korean-English bilinguals: the effect of age of arrival in the U.S. *Journal of the Acoustical Society of America*, 102(A).

Young, E. C. & Gilbert, H. R. (1988). An analysis of stops produced by normal children and children who exhibit velar fronting. *Journal of Phonetics*, **16**, 243-246.

APPENDIX

Production elicitation data:

APPENDIX A: English bisyllabic words with stress on the first syllable

	/● /	/● /	/🍁/
/● / or ■□/	Poodle Poodle Poodle	Tooler Tulip Tooling	Cooper Cooker Cooking
/● /	Potion Poker Potent	Total Totem Token	Coaching Cocoa Coco
/♀/	Potter Pocket Ponder	Tonic Toddler Topper	Cottage Copper Coffee

APPENDIX B. 27 Brazilian-Portuguese bisyllabic words with stress on the first syllable

	/● /	/● /	/🍁/
/● /	Pulo Puto Puro	Tunel Tudo Tussa	Cunha Cume Cujo

/● /	Poquer Poço Poça	Topo Toco Topo	Coxa Coco Cocho
/♀/	Pato Passo Passe	Taco Talo Tapa	Caça Casa Capa

APPENDIX C. Distracters: 22 English bisyllabic words with stress on the first syllable

Mary	Spinach
Searching	Mammal
Data	Rebel
Nipple	Forty
Window	Panic
Luggage	Foreign
Weapon	Cattle
Follow	Increase
Study	Himself
Lizard	Master
Growing	Sleepless

APPENDIX D. 22 Brazilian-Portuguese bisyllabic words with stress on the first syllable

Lingua	Pano
Minto	Nosso
Entre	Nome
Letra	Triste
Finco	Homem
Arte	Visto
Luta	Faça
Pomba	Pasta
Sogra	Cama
Pelo	Limpo
Mosca	Mora

APPENDIX E. Perception test answer sheet

Perception Answer Sheet

Identify English /● /

English Portuguese

1.

□● □ □ʌ□

□● □ □ʌ□

2.

□● □ □ʌ□

□● □ □ʌ□

3.

□● □ □ʌ□

□● □ □ʌ□

and so on until 120

Identify English /ʌ/

English Portuguese

1. □● □ □ʌ□

□● □ □ʌ□

2. □● □ □ʌ□

□● □ □ʌ□

3. □● □ □ʌ□

□● □ □ʌ□

and so on until 120

APPENDIX F. Sample questions of the production experiment

(1) Participants were instructed to translate the words elicited by the researcher's carrier sentence, for example:

"How would you say "cobre" (the metal, not the verb) in English?"

(2) Complete the sentence the researcher begins whenever they judged necessary, and that differed in many ways, for example:

1. A hot drink, generally black...
2. The main ingredient of chocolate is...

3. A synonym of hut or cabin...also a kind of cheese...
4. What is the continuous form of the act of ‘prepare food’....
5. the part of a piece of chothing in which you put your hands or keep objects or money or keys...